Genetic diversity and population genetic structure of red deer (Cervus elaphus) in the Scottish mainland, inferred by microsatellite markers and mitochondrial DNA control region sequences

Si'lvia Pe'rez-Espona

(BSc.Hons. Universitat Autònoma de Barcelona)
(MSc. The University of Edinburgh)

(photo modified from the DCS 1999-2000 annual report)

This thesis is submitted for the degree of Doctor of Philosophy to the Science and Engineering College, School of Biological Sciences The University of Edinburgh

July 2006

To my grandfather, Damián Pérez Blanca

Abstract

The red deer (*Cervus elaphus*) is the largest native terrestrial mammal in Britain, with most populations occurring in Scotland, where in fact the largest population of European red deer is found. Although extensive research on red deer genetics has been conducted in the long-term study area of the North Block in the island of Rum (Scotland), research on the genetics of red deer in mainland Scotland, which represent a more natural population of red deer, is limited. Studies on the genetic diversity and population structure of red deer on the Scottish mainland will provide important insights into the evolution and ecology of this species which will prove valuable for guiding future management policies.

To this end, in this study, I assessed the genetic diversity and population structure of red deer using 21 bi-parentally inherited microsatellite loci and 821bp length sequences of the mtDNA control region from 695 red deer individuals collected in the Scottish Highlands. Additionally, I took a landscape genetics approach and using Geographical Information System (GIS) techniques assessed the effect of natural and man-made features on contemporary gene flow of red deer across mainland Scotland. Sex-specific dispersal patterns were determined using three approaches: comparing microsatellite data between post-dispersal male and female red deer individuals; comparing population structure measures obtained by microsatellites and sequences of the mtDNA control region and by comparing mtDNA control region population differentiation measures between sexes. Finally, I conducted phylogenetic analyses to assess the extent of hybridisations and translocations of foreign deer stocks in mainland Scotland.

Despite the relatively small scale of the study area and the high dispersal capabilities of red deer, red deer on the Scottish mainland presented high levels of genetic diversity and significant population structure for both genetic markers, microsatellites (F_{ST} = 0.019; G_{ST}' = 0.084) and mtDNA control region sequences (Φ_{ST} = 0.3483). The landscape genetics approach indicated that landscape features play an important role on contemporary gene flow of red deer on the mainland, with sea lochs, roads, mountain slopes and forests located along the Great Glen valley being responsible for most of the genetic differentiation in the study area.

Sex-biased dispersal analyses, conducted using both genetic markers (microsatellites and mtDNA), revealed that male-biased dispersal was weak in our study area with male movements probably being predominant at a local scale (between neighbouring estates). In contrast, rarer long distance dispersal events which are more likely to be linked with colonisation of new areas were suggested to be predominantly female-biased.

In terms of management, results from this study suggest that past management practices have not strongly affected the genetic integrity, genetic diversity and population genetic structure of red deer on the Scottish mainland. Phylogenetic analyses revealed that none of the red deer individuals included in this study had introgressed mtDNA from other species or subspecies despite the numerous introductions of foreign species of deer in the Scottish mainland such as wapiti (*Cervus canadensis*) or sika deer (*Cervus nippon*). Furthermore, only few localised individuals were found to have potentially descended from translocation events. Results from this study also support the continuation of current policies for the management of red deer on mainland Scotland by delimiting Deer Management Groups (discrete populations or herds of deer) according to natural and man-made landscape features.

Acknowledgements

First of all I would like to thank all the deer stalkers and managers from the estates of Ardgour, Ardverikie, Ben Alder, Clunes, Conaglen, Corrour, Forest Lodge, Glencoe, Glencreran, Glenkinglas, Glenstrae, Mamore, Kintail and South Affric for collecting all samples used in this study, their hard work and their professionalism ensured an invaluable collection of samples and therefore they are partly responsible for the success of this project. The Macaulay Development Trust funded this project and supported me with a Ph.D. studentship. The project is part of the research programme on population performance and distribution of Scottish red deer carried out by Javier Pérez-Barbería in the Ecology of Grazed Ecosystems department in the Macaulay Institute. I thank my supervisors Javier Pérez-Barbería, Josephine Pemberton, Iain Gordon and Chris Jiggins for fruitful discussions during the course of the Ph.D. and for all their commitment on reading all the manuscripts presented in this thesis. In particular, I would like to thank Javier Pérez-Barbería for giving 100% to the project and always being there, in the good and the bad moments.

From the Macaulay Institute I would like to thank Jim McLeod for doing and excellent job constructing the GIS matrices and landscape maps, Russell Hooper the GIS guru for providing some of the maps presented in this thesis and Angela Sibbald for providing information from unpublished studies on GPS-collared red deer in mainland Scotland. Steve Albon is also thanked for dealing with administrative issues at the Macaulay Institute. At the University of Edinburgh I would like to thank Andy Gillies, Jill Lowe, Anna Montazam and Jenna Paxton for all their hard work running the Institute of Evolutionary Biology sequencing services. Dan Nussey is thanked for helpful discussions about red deer. Margarita Beltran, Angeles de Cara, Barbara Craig, Felicity Jones, Martim Melo, Rita Covas, Roberta, Kati, Helen Senn, Lel, Jun Wang, Bengt Hanson such good colleagues, and my dear friend Dario Beraldi for sharing all the ups and downs during the Ph.D.

There are lot of people who I met in Scotland during the past years along the several jobs and studies I undertook in Scotland among many people I would like to thank Gloria Bueno, Ana Hickey, Pablo Sanz, Eileen Sanz, Susana Favela, Mariana Gabarrot, Chris Kettle, Wendy Martin, Adriana Otero-Arnaiz and Pablo Fuentes for all their support and friendship. My friends from the Universitat Autònoma de Barcelona Elisenda Pastó, Julita Ocaña, Olga Montanya, Jordi Pascual, Clara Pons, Montse Masià, Flors Moreno, Sam Matellán and Enric Segalés are thanked for always supporting me despite the distance. I also thank Jesús Mávarez for invaluable advice and fruitful discussion on population genetics.

My family in particular my parents Juan José Pérez Cámara and Maria Rosa Espona i Casals, my sister Mari and her husband Sergio Martínez and my little nephew Ivan, are thanked for their continuous support and love. Last but not least, I would like to thank Will Goodall-Copestake for all his advice on molecular biology and phylogenetics and for all the good moments during the past years.

Table of contents

Chapter 1

General Introduction

1.1 Red deer (*Cervus elaphus*)

The red deer (*Cervus elaphus*) is one of the most widespread and best studied deer species in the world (Clutton-Brock & Albon 1989; Ludt *et al.* 2004; Fig. 1.1). Red deer in Europe have been strongly influenced by man either indirectly through habitat destruction and grazing competition with livestock or directly through selective hunting and translocations of foreign stocks in order to improve trophy quality (Lowe & Gardiner 1974; Herzog *et al.* 1991; Hartl *et al.* 2003; Feulner *et al.* 2004).

Figure 1.1. Scottish red deer. Painting "Monarch of the Glen", c. 1851 by Sir Edwin Landseer R.A.

The largest population of red deer in Europe is found in Britain, which accounts for 30% of the total population (Clutton-Brock & Albon 1989). The red deer is the largest terrestrial mammal in Britain and together with roe deer (*Capreolus capreolus*), the only native deer (Clutton-Brock & Albon 1989). However, the current distribution of red deer in Britain is very patchy, with most of the populations occurring in Scotland and only few small and scattered populations being found in England (Whitehead 1964; Ward 2005; Fig. 1.2). In Scotland, red deer are distributed in the islands, in SW Scotland and in the Highlands where the highest densities of red deer are found, with an estimated population of about 400,000 individuals thought to occupy an area of 300,000 km^2 (Clutton-Brock & Albon 1989; Clutton-Brock *et al.* 2004).

Figure 1.2. Distribution of red deer in Britain. Blue represents current distribution and green the areas where red deer once occurred (map taken from The Deer Initiative website - http://wwww.thedeerinititiave.co.uk)

The first appearance of *C. elaphus* was in the Middle Pleistocene Cromerian interglacial of Europe (c. 400,000 years BP) (Lister 1984). The red deer is thought to have continuosly occurred in Britain throughout the last (Ipswichian) interglacial period (c. 125,000-110,000 years BP) and had a continuous presence in Scotland since the end of the last glaciations (c. 11,000 years BP) (Lister 1984) when forests with interspersed open areas were abundant throughout Scotland (Clutton-Brock & Albon 1989). Although red deer evolved as a species occupying low-ground habitats such as heaths, forests or forests edges, most red deer in Scotland had to adapt to live in open hill areas due to the disappearance of forests starting with the development of farming cultures (c. 5,000 years BP) and continuing until most of the 18th century (Whitehead 1964; Lister 1984; Clutton-Brock & Albon 1989). As a result of the harsher environmental conditions in open hill habitats, including limited shelter and poor quality grazing, red deer in open hill areas have smaller body size than red deer living in forests or deer farms (Clutton-Brock & Albon 1989).

1.2 Red deer management in Scotland

Due to an increase in red deer population sizes during the past decades, probably as a result of combination of a series of mild winters and a decreased competition for grazing with sheep (Clutton-Brock & Albon 1992), many Scottish red deer populations are currently managed in order to control population growth and their impact on habitats of conservation value, forestry and agriculture. Red deer are also managed to provide trophy hunting in sporting estates by traditional deer stalking (Clutton-Brock *et al.* 2004; Deer Commission for Scotland website – http://www.dcs.gov.uk). The Deer Commission for Scotland (DCS) is the Government agency responsible for regulating the conservation and sustainable management of deer in Scotland (Deer Commission for Scotland website). Red deer, as wild species, are not owned by anybody but landowners are entitled to capture and kill them if present on their land; however, landowners are also responsible for the welfare and for the control of deer numbers (Clutton-Brock & Albon 1989; Association of Deer Management Groups website -http://www.deer-management.co.uk). The range of red deer in Scotland is divided into Deer

Management Groups (DMGs) which consist of groups of estates or other landholdings covering an area delimited by natural or man-made features that are thought to contain distinct populations or herds of deer (Clutton-Brock & Albon 1989; Asociation of Deer Management Groups -ADMG website; Fig. 1.3). However, DMGs are formed voluntarily between estates in order to collaborate in the management of red deer and are not stable divisions, sometimes being re-arranged for practical or political reasons, with the disappearance, merging or division of group areas not necessarily reflecting natural boundaries of populations (Iain Gordon, personal communication).

Figure 1.3. Distribution of DMGs covering most of the distribution of red deer in Scotland

The day-to-day management of red deer is carried out by individual estates or other landholdings. Professional and experienced deer stalkers carry out regular counts of the populations, culling (hunting) operations, research and survey work (Clutton-Brock & Albon 1989; DCS and ADMG websites). Most of the estates are privately owned and provide fishing, shooting and deer stalking for sport to private lets, but can also be involved in other activities such as farming, forestry and tourism (Bullock *et al.* 1998; ADMG website). However, some other estates are owned by organisations such as The National Trust for Scotland, The John Muir Trust, The Royal Society for the Protection of Birds, The Scottish Wildlife Trust and the Forestry Commission where red deer management is conducted in order to protect habitats of conservation value or forestry (Clutton-Brock *et al.* 2004; ADMG website). Legal culling seasons for red deer in Scotland are from 1st July-20th October for males and from 21th October-15th February for females (Clutton-Brock & Albon 1989; DCS website).

The management of red deer on the Scottish mainland is complicated by many factors, one of which is that little is known about the scale over which red deer disperse and what factors affect dispersal. In particular, we can ask whether the Deer Management Groups, identified by the Deer Commission for Scotland, are likely to be managing real subpopulations of deer or not. Due to the importance of red deer in the rural Scottish economy, with the current value of deer lets in Scotland estimated at between £8-10 million/year and venison sales about £5 million/year (Clutton-Brock & Albon 1989; Scottish Natural Heritage 1994; ADMG website), debates have arisen recently between estate managers and landowners about the movement of red deer between estates, generating a need for dispersal studies on this species.

1.3 Field versus molecular-based studies of dispersal

Field-based studies such as mark-release-recapture and radio- or satellite-tracking can provide direct measures of dispersal (i.e. movement of individuals) which might be important for local and immediate management actions. However, field-based

methods are logistically difficult for studies of large mammals with high dispersal capabilities and only a small number of individuals can be followed due to the associated costs with these techniques (Slatkin 1985; Koening *et al.* 1996). Moreover, rarer long distance movements of individuals are less likely to be detected with field-based methods and the information obtained from field-based methods might only be valid for the temporal scale of the duration of the study (Koening *et al.* 1996).

The use of molecular markers to infer dispersal has become increasingly popular as they can overcome some of the difficulties associated with field-based studies (Slatkin 1985; Paetkau 1995; Neigel 1997) and can sometimes reveal a different aspect of dispersal from that suggested by field-based methods (e.g. Jones 1987; Dobson 1994). In addition, estimates obtained from molecular markers such as nuclear microsatellite loci and mitochondrial DNA (mtDNA) sequences can provide insights into gene flow and population history (Sunnucks 2000) which are important in order to define appropriate scales for long-term conservation and management programmes (Avise 1994; Moritz 1994; Humphries *et al.* 1995).

1.4 The effect of landscape features on dispersal

Landscape heterogeneity can strongly affect dispersal patterns and together with the dispersal capabilities and the social structure of an organism, will determine the distribution of genetic diversity (Gustafson & Gardner 1996; With *et al.* 1997). The effect of landscape features on gene flow is still poorly understood as the identification of which features affect the gene flow of a species are often difficult to identify due to the complex mixture of land cover types that make up a landscape (Manel *et al.* 2003). However, new developments in the area of landscape genetics, a discipline which combines landscape ecology and population genetics, are providing new insights on how landscape features affect the genetic structure of populations (Manel *et al.* 2003).

1.5 Social structure and sex-biased dispersal of red deer

Like many other polygynous mammals, the social structure of red deer is characterised by strong female philopatry and male-biased dispersal (Greenwood 1980; Clutton-Brock & Albon 1989). Studies conducted in the North Block study area of the Island of Rum (off the west coast of Scotland) where deer are individually monitored, have shown that male and female red deer generally segregate for most of the year, occupying geographically separate areas and only coming together when the mating season (rut) starts in late September (Clutton-Brock *et al.* 1982; Clutton-Brock & Albon 1989; Conradt *et al.* 2000, 2001). Red deer females are strongly philopatric, generally occupying the same particular area throughout their lives and form matrilineal groups consisting of different generations of related females with shared home ranges (Clutton-Brock *et al.* 1982; Clutton-Brock & Albon 1989; Albon *et al.* 1992). In contrast, males generally disperse from the mother's group when they are 2 to 4 years old (natal dispersal; Clutton-Brock *et al.* 1982; Clutton-Brock & Albon 1989; Clutton-Brock *et al.* 2002; Catchpole *et al.* 2004). Once juvenile stags abandon the female group they can live as solitary individuals or often they join a group of other males (bachelor groups, Fig. 1.4; Clutton-Brock *et al.* 1982; Clutton-Brock & Albon 1989).

Fig. 1.4. Bachelor group of male red deer (photo taken from ArKive website, http://www.arkive.com)

In early September, stags disperse temporarily from their summer range and search for areas with females where they stay during the rut (Clutton-Brock *et al.* 1982; Clutton-Brock & Albon 1989). During the rut (October to November) adult stags compete with each other by means of roaring contests and fights in order to take possession of and defend harems of hinds (Clutton-Brock *et al.* 1982; Clutton-Brock & Albon 1989). Red deer are considered to be strongly polygynous, with dominant males siring all or most of the progeny (Pemberton *et al.* 1992). From late May to early July females give birth generally to a single calf, very rarely two (Clutton-Brock *et al.* 1982; Clutton-Brock & Albon 1989).

The population history, environmental conditions and management practices of red deer in mainland Scotland differ substantially from those found in the North Block area in Rum, and therefore the costs and the benefits of sex-specific dispersal might also vary between the mainland and island populations (Boujemadi *et al.* 1999; Whitlock 2001; Wiens 2001). Discrepancies in the direction and extent of sex-biased dispersal between populations of the same species have already been found in other mammals such as Eurasian badger (*Meles meles*; Pope *et al.* 2006) and the European rabbit (*Oryctolagus cuniculus*; Richardson *et al.* 2002).

Knowledge of sex-specific dispersal patterns of red deer will also be valuable for conservation and management programs in the mainland of Scotland as different management actions might be required for each sex (e.g. Maehr *et al.* 2002; Zenger *et al.* 2003; Rubin & Bleich 2005).

1.6 Objectives of the study

In this study, I used two widely used molecular markers in population biology, microsatellite markers and mtDNA control region sequences to:

1) Assess the genetic diversity and population structure of red deer in mainland Scotland.

2) Assess the effect of landscape features on population genetic structure of red deer in the Scottish mainland.

3) Assess the extent to which sex-specific dispersal patterns are influencing the distribution of genetic diversity of red deer in the Scottish mainland.

4) Assess the extent to which past management practices have affected the genetic diversity and population structure of red deer in the Scottish mainland.

5) Discuss the implications of the results from this study for red deer management in the Scottish mainland.

1.7 Thesis outline

A wide range of microsatellite markers is available for use in population genetics of red deer due to the success of cross-amplification of microsatellite loci isolated from bovids in cervid species (Slate *et al.* 2002). However, amplification of microsatellite loci isolated from a particular species might sometimes not produce reliable genotyping when cross-amplified in other species. Optimisation procedures such as re-designing original primers and modifying polymerase chain reaction (PCR) conditions often yield improvements in amplification success and increase the quality of genotyping. Furthermore, for population genetic studies in which a large number of individuals needs to be genotyped for several loci in order to obtain reliable estimates of population parameters, multiplexing techniques (the amplification of several loci in a single PCR and the loading of several PCR products in a single lane of a sequencer) can aid the development of reliable and efficient

high-throughput genotyping protocols in order to make studies time and cost effective (e.g. Thomas *et al.* 1999; Cryer *et al.* 2005). In Chapter 2, I described an efficient and reliable microsatellite multiplex high-throughput protocol which I developed by re-designing primers and optimising PCR conditions for 21 microsatellite markers previously isolated from red deer, wapiti, reindeer, sheep and cattle. Descriptive analyses of the performance of the microsatellite multiplex protocol for population genetic studies and other applications such as parentage and forensic analyses are also given in Chapter 2.

Natural and man-made landscape features have been the main variables used to define Deer Management Groups in Scotland (Clutton-Brock & Albon 1989; Association of Deer Management Groups website). However, for organisms with high dispersal capabilities such as red deer, geographical or political boundaries might not always reflect boundaries of natural populations (Cegelski *et al.* 2003). Using the microsatellite multiplex protocol developed in Chapter 2, I genotyped a large number of red deer individuals (n = 695) collected from an E-W transect across the Highlands. The study area presented considerable landscape heterogeneity, with several natural and man-made features that could potentially affect the dispersal of red deer. In Chapter 3, following a landscape genetics approach and using Geographical Information Systems (GIS) techniques I determined the extent to which several natural and man-made landscape features (sea lochs, inland lochs, forests, roads, mountain slopes, rivers and railways) influence the population structure of red deer in mainland Scotland. Results from this chapter show that despite the high dispersal capabilities of red deer significant population genetic structure is found in the mainland of Scotland and that the major differentiation between populations is due to several landscape features acting as gene flow barriers along the Great Glen (in Scotland a valley is a glen).

In Chapter 4, I assessed the direction and extent of current sex-biased dispersal of red deer in the Scottish Highlands. Sex-biased dispersal tests were undertaken by comparing estimates of population structure, relatedness and assignment tests

between male and female post-dispersal red deer (4 years or older) (Goudet *et al.* 2002). In addition, I conducted individual-based spatial autocorrelation analyses to assess the scale at which any possible bias in dispersal was detectable. Male-biased dispersal was significant for comparisons of population structure, although the bias was not as strong as expected for a polygynous mammal. Furthermore, spatial autocorrelation analyses revealed similar trends of kinship and geographical distance within each sex and only suggested weak male-biased dispersal at small distance intervals that might represent greater movement of males than females between neighbouring estates.

Although microsatellite loci are the marker of choice for population genetic studies (Sunnucks 2000), it is important to acquire independent information from other genetic markers that cover different evolutionary time scales in order to provide helpful guidelines for management (Crandall *et al.* 2000). Mitochondrial DNA markers such as the control region (mtDNA CR) have been widely used to infer population history as they can provide a more historical perspective on the evolution and demography of populations than microsatellite markers (Avise 1994; Sunnucks 2000). For the last data chapter in this thesis (Chapter 5), I analysed sequences of the mtDNA CR for the same 695 red deer previously genotyped at 21 microsatellite markers in order to infer historical population genetic structure and to assess the extent to which past management practices have affected the native gene pools of Scottish Highland red deer.

Due to the maternal inheritance of mtDNA in mammals, mtDNA patterns of population genetic structure reflect female dispersal and therefore comparisons of patterns of population structure obtained with mtDNA with those obtained with the bi-parentally inherited microsatellite markers can be used to assess sex-biased dispersal (Ennos 1994; Fitzsimmons *et al.* 1997; Prugnolle & De Meeus 2002). In Chapter 5, I re-assess sex-biased dispersal by comparing estimates of population structure obtained with microsatellite and mtDNA sequences.

The analyses conducted in Chapter 5 revealed that the influence of man in Scottish mainland red deer populations has not strongly affected the genetic diversity and population structure of red deer. Most of the study populations presented high levels of genetic diversity and none of the individuals included in this study presented introgressed mtDNA haplotypes of sika, wapiti or other introduced species or subspecies of deer. Furthermore, significant mtDNA population structure was found in the study area with only a few localised individuals suggesting possible translocation events. Haplotype relationships and demographic analyses using mismatch distribution and neutrality tests suggested that the Scottish mainland red deer population has originated from a population expansion which probably occurred during the Pleistocene.

Chapters 2-5 are presented in manuscript format with full literature reviews of the subject area in their introductions. The main conclusions from this study, possible guidelines for current and future management of Scottish Highland red deer and future research that might develop from results obtained from this study are described in the last chapter of this thesis, Chapter 6.

1.8 References

Albon SD, Staines HJ, Guinness FE, Clutton-Brock TH (1992) Density-dependent changes in the spacing behaviour of female kin in red deer. *Journal of Animal Ecology* **61**, 131-137.

Association of Deer Management Groups website. http://deer-management.co.uk.

Avise JC (1994) *Molecular Markers, Natural History and Evolution* Chapman & Hall, New York.

Boudjemadi K, Lecomte J, Clobert J (1999) Influence of connectivity on demography and dispersal in two contrasting habitats: an experimental approach. *Journal of Animal Ecology* **68**, 1207-1224.

Bullock CH, Elston DA, Chalmers NA (1998) An application of economic choice experiments to a traditional land use-deer hunting and landscape change in the Scottish Highlands. *Journal of Environmental Management* **52**, 335-351.

Catchpole EA, Fan Y, Morgan BJT, Clutton-Brock TH, Coulson T (2004) Sexual dimorphism, survival and dispersal in red deer. *Journal of Agricultural, Biological and Environmental Statistics* **9**, 1–26.

Cegelski CC, Waits LP, Anderson NJ (2003) Assessing population structure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. *Molecular Ecology* 2003 **12**, 2907-2918.

Clutton-Brock TH, Guinness FE, Albon SD (1982) *Behaviour and Ecology of Two Sexes* University of Chicago Press, Chicago.

Clutton-Brock TH, Albon SD (1989) *Red deer in the Highlands* BSP Professional Books, Oxford.

Clutton-Brock TH, Albon SD (1992) Trial and error in the Highlands. *Nature* **358**, 11-12.

Clutton-Brock TH, Coulson T, Milner JM (2004) Red deer stocks in the Highlands of Scotland. *Nature* **429**, 261-262.

Conradt L, Clutton-Brock TH, Guinness FE (2000) Sex differences in weather sensitivity can cause habitat segregation: red deer as an example. *Animal Behaviour* **59**, 1049-1060.

Conradt L, Gordon IJ, Clutton-Brock TH, Thomson D, Guinness FE (2001) Could the indirect competition hypothesis explain inter-sexual site segregation in red deer (*Cervus elaphus* L.)?. *Journal of Zoology* **254**, 185-193.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution* **15**, 290-295.

Cryer NC, Butler DR, Wilkinson MJ (2005) High throughput, high resolution selection of polymorphic microsatellite loci for multiplex analysis. *Plant Methods* **1**.

Deer Commission for Scotland website. http://www.dcs.gov.uk.

Dobson FS (1994) Measures of gene flow in the Columbian ground squirrel. *Oecologia* **100**, 190–195.

Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* **72**, 250-259.

Feulner PGD, Bielfeldt W, Zachos FE (2004) Mitochondrial DNA and microsatellite analyses of the genetic status of the presumed subspecies *Cervus elaphus montanus* (Carpathian red deer). *Heredity* **93**, 299-306.

Fitzsimmons NN, Moritz C, Limpus CJ, Pope L, Prince R (1997) Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. *Genetics* **147**, 1843-1854.

Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology* **11**, 1103-1114.

Greenwood PJ (1980) Mating system, philopatry, and dispersal in birds and mammals. *Animal Behaviour* **28**, 1140-1162.

Gustafson EJ, Gardner RH (1996) The effect of landscape heterogeneity on the probability of patch colonization. *Ecology* **77**, 94-107.

Hartl G, Zachos FE, Nadlinger K, *et al.* (2005) Allozyme and mitochondrial DNA analysis of French red deer (*Cervus elaphus*) populations: genetic structure and its implications for management and conservation. *Mammalian Biology* **70**, 24-34.

Herzog S, Mushövel C, Hattemer HH, Herzog A (1991) Transferrin polymorphism and genetic differentiation in *Cervus elaphus* L. (European red deer) populations. *Heredity* **67**, 231-239.

Humphries CJ, Williams PH, Vane-Wright RI (1995) Measuring biodiversity value for conservation. *Annual Review of Ecology and Systematics* **26**, 93–111.

Jones WT (1987) Dispersal patterns in kangaroo rats (*Dipodomys spectabilis*). In: *Mammalian Dispersal Patterns: The Effects of Social Structure on Population Genetics* (eds. Chepko-Sade BD, Halpin ZT), p. 119–127. University of Chicago Press, Chicago.

Koening WD, Van Vuren D, Hooge PN (1996) Detectability, philopatry, and the distribution of dispersal distances in vertebrates. *Trends in Ecology and Evolution* **11**, 514-517.

Lister AM (1984) Evolutionary and ecological origins of British deer. *Proceedings of the Royal Society of Edinburgh* **82B**, 205-299.

Lowe VPW, Gardiner AS (1974) A re-examination of the subspecies of Red deer (*Cervus elaphus*) with particular reference to the stocks in Britain. *Journal of Zoology (London)* **174**, 185-201.

Ludt CJ, Schroeder W, Rottman O, Kuehn R (2004) Mitochondrial DNA phylogeography of red deer (*Cervus elaphus*). *Molecular Phylogenetics and Evolution* **3**, 1064-1083.

Maehr DS, Land ED, Shindle DB, Bass OL, Hoctor TS (2002) Florida panther dispersal and conservation. *Biological Conservation* **106**, 187-197.

Manel S, Schwartz M, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution* **18**, 189-197.

Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution* **9**, 373-375.

Neigel JE (1997) A comparison of alternative strategies for estimating gene flow from genetic markers. *Annual Review of Ecology & Systematics* **28**, 105-128.

Paetkau D (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* **4**, 347-354.

Pemberton JM, Albon SD, Guinness FE, Clutton-Brock TH, Dover GA (1992) Behavioural estimates of male mating success tested by DNA fingerprinting in a polygynous mammal. *Behavioral Ecology* **3**, 66-75.

Pope LC, Doming-Roura X, Erven K, Burke T (2006) Isolation by distance and gene flow of the Eurasian badger (*Meles meles*) at both local and broad scale. *Molecular Ecology* **15**, 371-386.

Prugnolle F, De Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity* **88**, 161-165.

Richardson BJ, Hayes RA, Wheeler SH, Yardin MR (2002) Social structure, genetic structures and dispersal strategies in Australian rabbit (*Oryctolagus cuniculus*) populations. *Behavioral Ecology and Sociobiology* **51**, 113-121.

Rubin ES, Bleich VC (2005) Sexual segregation: a necessary consideration in wildlife conservation. In: Sexual segregation in vertebrates: ecology of the two sexes (eds. Ruckstuhl KE, Neuhaus P), p. 379–391. Cambridge University Press, Cambridge.

Scottish Natural Heritage (1994) Red deer and the natural heritage. SNH policy paper. Scottish Natural Heritage, Battleby.

Slate J, Van Stijn TC, Anderson RM, *et al.* (2002) Genetic linkage map and the evolution of ruminant genomes. *Genetics* **160**, 1587-1597.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology & Systematics* **16**, 393-430.

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology and Evolution* **15**, 199-203.

Thomas MG, Bradman N, Flinn HM (1999) High throughput analysis of 10 microsatellite and 11 diallelic polymorphisms on the human Y-chromosome. *Human Genetics* **105**, 577-581.

Ward AI (2005) Expanding ranges of wild and feral deer in Great Britain. *Mammal Review* **35**, 165-173.

Whitehead GK (1964) *The deer of Great Britain and Ireland* Routledge & Kegan Paul, London.

Whitlock MC (2001) Dispersal and the genetic properties of metapopulations. In: Dispersal (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), p. 273–282. Oxford University Press, Oxford.

Wiens JA (2001) The landscape context of dispersal. In: Dispersal (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), p. 96-109. Oxford University Press, Oxford

With KA, Gardner RH, Turner MG (1997) Landscape connectivity and population distributions in heterogeneous environments. *Oikos* **78**, 151-159.

Zenger KR, Eldridge MDB, Cooper DW (2003) Intraspecific variation, sex-biased dispersal and phylogeography of the eastern grey kangaroo (*Macropus giganteus*). *Heredity* **91**, 153-162.

Chapter 2

An efficient and reliable 21 microsatellite multiplex protocol for high throughput population genetic studies of red deer (*Cervus elaphus*)

S. Pérez-Espona,[*§]

*Institute of Evolutionary Biology, School of Biological Sciences,The University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

§ The Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK

Corresponding author:

Sílvia Pérez-Espona,

Institute of Evolutionary Biology, School of Biological Sciences, The University of Edinburgh

The King's Buildings, West Mains Road, Edinburgh EH9 3J, UK

Fax: ++ 44 (0)131 650 6564

Tel: ++44 (0)131 651 3612

E-mail: s.perez-espona@sms.ed.ac.uk

2.1 Abstract

The red deer (*Cervus elaphus*) is one of the most importantally socioeconomic wild mammal species in Europe and as such its population structure and distribution has long been influenced by humans. In order to study the genetic diversity and population structure of red deer on the mainland of Scotland an efficient and reliable high-throughput microsatellite genotyping protocol was designed by optimising microsatellite markers already isolated for cattle (*Bos taurus*), sheep (*Ovis aries*), reindeer (*Rangifer tarandus*), wapiti (*Cervus elaphus canadensis*) and red deer. Nineteen microsatellite loci were amplified in five multiplex polymerase chain reactions (PCRs) and two loci were amplified in uniplex. The 21 loci were grouped into three multiloading panels. The multiplex kit was successfully used to genotype 871 samples from red deer collected across the Scottish Highlands. These loci presented moderate to high polymorphism (7-30 alleles), and a mean expected heterozygosity (H_E) of 0.804. Parental probability of exclusion were $P_{E1} = 0.9999$ and $P_{E2} = 1$, and the probability of identity $P_{ID} = 2.63 \times 10^{-27}$. The multiplex kit described here is an efficient and reliable tool for population studies of red deer and for identification of illegal trafficking of trophies.

Keywords *Cervus elaphus*, genetics, microsatellites, multixing, multiplex PCR, population, probability of exclusion, probability of identity, red deer.

2.2 Introduction

The red deer is widely distributed from Europe across Asia westward to North America (Clutton-Brock *et al.* 1982). Red deer are classified into a number of subspecies and populations across their range, and occur as wild animals, semi-domestic animals living in parks and as farmed animals for the production of venison (Geist *et al.* 1999). The population structure of red deer in Europe reflects its history as an important socioeconomic species that has been intensively managed over many centuries, including numerous translocations of animals between countries in order to improve hunting trophy quality (Lowe & Gardiner 1974; Hartl *et al.* 2003). Additionally, the population structure has been affected by anthropogenic changes to the landscape structure and habitat distribution because of urbanisation, roads, railway lines, forest clearing, planting and agricultural development (e.g. Hartl *et al.* 1990; Ströhlein *et al.* 1993; Martínez *et al.* 2002). In order to draw up effective management or conservation policies it is important to identify management units (*sensu* Moritz 1994 , i.e. populations with significantly different genotypes), assess population genetic diversity and quantify gene flow between different populations.

Microsatellites have become the marker of choice in many population genetic studies due to their ubiquity in eukaryotic genomes, generally high levels of polymorphism, codominant inheritance and often easy and reliable scoring (Jarne & Lagoda 1996; Luikart & England 1999b; Schlötterer 2002). Species-specific primers for microsatellites can be developed by using cloning techniques; however, cloning often requires a substantial amount of time and effort (Hammond *et al.* 1998). A valuable characteristic of microsatellites is that cross-amplification between closely related species is often possible due to conserved microsatellite flanking regions. In ungulates in particular, the transfer of microsatellite primer sets between related species has been successfully applied (e.g. Moore *et al.* 1991; Engel *et al.* 1996; Kuehn *et al.* 1996; Røed 1998; Slate *et al.* 1998; Luikart *et al.* 1999a; Bonnet *et al.*, 2002; Galan *et al.* 2003). Furthermore, thousands of genetic markers have become available with the publication of genome mapping projects for some economically important bovid species, thus greatly increasing the potential number of markers for

use in cervids (e.g. Barendse *et al.* 1994; Bishop *et al.* 1994; Maddox *et al.* 2001; Slate *et al.* 2002).

For population studies, where a large number of individuals are required to be genotyped at several microsatellite loci, an efficient high-throughput protocol is necessary for researchers to be able to improve the time and economic cost-effectiveness of analyses (e.g. Thomas *et al.* 1999; Cryer *et al.* 2005). With recent developments in PCR technology and improvement in the quality of many chemical reagents, multiplex PCR (i.e. the amplification of several loci in a single reaction) allows the genotyping of a large number of individuals across many loci (Chamberlain *et al.* 1988; Ziegle *et al.* 1992; Henegariu *et al.* 1997). In addition, variation in microsatellite fragment size and the availability of different fluorescent dyes with which to label primers also allow the loading of several microsatellite sets in a single sequencer lane (multimixing) (Ciofi *et al.* 1998). Multimixing panels have already been applied in deer studies, for example Bonnet *et al.* (2002) developed a multiplex panel consisting of 11 microsatellite markers to study population structure and diversity in four tropical *Cervus* species, and Galan *et al.* (2003) developed a 12 microsatellite loci multiplex panel for population studies of roe deer (*Capreolus capreolus*). Although several studies have described sets of microsatellite loci for genetic studies of red deer and have exploited the advantages of multimixing (e.g. Slate *et al.* 1998, Haanes *et al.* 2005) only one has exploited the advantages of PCR multiplexing (Niemczewski *et al.* 2002). However, from their 27 loci multiplex kit, 15 loci did not amplify, 4 were monomorphic and 5 had very low heterozygosity and polymorphic information content. Here is described an efficient and reliable high-throughput multiplex kit consisting of 21 microsatellite loci designed for studies of several aspects of population structure and gene flow in red deer on the mainland of Scotland.

2.3 Materials and Methods

2.3.1 Sample collection and DNA extraction

Red deer samples from 871 wild individuals (400 males, 299 females and 172 foetuses), legally shot in hunting and culling operations carried out in open hill habitats in the Scottish Highlands, were collected by stalkers or estate factors during the 2003-2004 hunting season. Samples consisted of a jaw, an ear or a whole foetus and they were part of a more extensive study on the species in Scotland. Tissue samples were stored either in a -20ºC freezer or in tubes containing 100% ethanol. Genomic DNA was extracted from muscle tissue using the DNAace Spin Tissue Mini Kit (Bioline) or with the DNEasy Tissue Kit™ (QIAGEN), following the manufacturer's instructions.

2.3.2 Microsatellite marker selection

Literature search was undertaken to select microsatellite markers already isolated in cattle (*Bos taurus*), sheep (*Ovis aries*), reindeer (*Rangifer tarandus*), wapiti (*Cervus elaphus canadensis*) and red deer (*Cervus elaphus*) that were known to amplify in red deer or closely related species. From these markers, 44 were chosen as they had shown high polymorphism and high heterozygosity in red deer or related species (Røed & Midthjell 1998; Slate *et al.* 1998; Goodman *et al.* 1999; Fickel & Reinsch 2000; Bonnet *et al.* 2002; Jones *et al.* 2002, Niemczewski *et al.* 2002; Galan *et al.* 2003; Vial *et al.* 2003). The location of the loci in deer, cattle and sheep chromosome maps was checked when available to ensure that loci were not linked by searching in the Arkdb database (http://www.thearkdb.org).

To assess the potential utility of these primers for multiplex PCR, variation in annealing temperature, GC content and primer-primer dimerisation was checked using the program Oligotoolkit (Operon, http://www.operon.com/oligos/toolkit.php). Most primer pairs were manually re-designed to achieve a predictive annealing temperature (Tm) between 58ºC and 65ºC (as high annealing temperatures decrease

the amplification of non-specific products), and to avoid, as much as possible, high predicted primer dimerisation. Those primer pairs following the above criteria were tested for amplification using 16 individuals sampled across the whole study area. A gradient PCR for annealing temperatures ranging from 55-65°C was performed for each primer set using identical PCR conditions in order to identify the optimal annealing temperature for multiplexing. Uniplex PCRs were carried out using 10-15 ng of DNA, 0.2mM dNTPs, 1x NH_4 Buffer, 3mM $MgCl_2$, 0.2mM of each primer (one of them fluorescently labelled with either 6-FAM, VIC, NED or PET, Applied Biosystems), 0.3625 units of BIOTAQ™ polymerase (Bioline) and double processed tissue culture distilled H_2O (Sigma) to bring the volume up to 15μl. The amplification products were diluted both at 1:25 and 1:100 to test for optimal loading concentration. 1μl of the diluted PCR product was mixed with 9 μl of loading mix (consisting of 1ml Hi-Di™ Formamide + 1μl of the internal size standard Genescan LIZ 500 ladder, Applied Biosystems), denaturated by incubation at 95 °C for 2 min. and run on a capillary ABI 3730 DNA Analyzer (Applied Biosystems). Fragment analysis was performed using the software GeneMapper™ version 3.0 (Applied Biosystems).

2.3.3 Multiplex PCR and multimixing

Those microsatellite loci that gave good quality amplification, were polymorphic and provided unambiguous scoring were selected for multiplex-PCR trials on the same 16 individuals used for uniplex PCR reactions. Primer pairs were grouped into sets and for those primer pairs with overlapping ranges a different fluorescent label was selected. Multiplex PCR trials were performed using different combinations of primer sets to achieve the maximum number of loci amplified in a single PCR that would yield optimal amplification and unambiguous scoring. Multiplex PCR reactions were performed using the QIAGEN® Multiplex PCR kit and consisted of 10-15ng of DNA, 7.5μl of QIAGEN multiplex PCR master mix (HotStar Taq® DNA Polymerase, multiplex PCR buffer containing 6mM $MgCl_2$ and dNTP mix), an initial primer concentration of 0.2μM, and RNase free distilled H_2O (QIAGEN) to make up a final volume of 15μl. The microsatellite cycling protocol involved an initial

activation step of 95°C for 5 min., a 3 step-cycling consisting of a denaturing step of 94°C for 30 sec., annealing at 59°C for 1min. 30sec., ramping at 0.3°C/sec. to an extension step of 72°C of 1min. 30sec. The cycle was repeated 29 times and was followed by a final extension of 60°C for 30 minutes. Trial multiplex PCR sets were again diluted either 1:25 or 1:100 and run in a capillary 3730 sequencer (Applied Biosystems). To ensure allele amplification consistency, results from co-amplifications were compared to those obtained when amplifying loci in uniplex. For all successful co-amplifications, primer concentrations were refined in order to have similar strength co-amplification of loci within the same multiplex PCR set. When multiplex-PCR sets were defined, multimixing loading trials were carried out to optimise multimixing sets, i.e. maximum number of multiplex PCR sets loaded in a single sequencer run without compromising allele scoring. Multimixing sets were diluted both at 1:25 and 1:100 and run in the capillary sequencer under the same conditions as for the uniplex amplified loci.

2.3.4 Measures to control for consistent genotyping

All 871 individuals sampled (males, females and foetuses) were genotyped for the microsatellites included in the multiplex kit. The following measures were taken to check for consistency of genotyping: 1) the same six to eight standard samples were always included in each of the sequencer runs, 2) foetus genotypes were always compared to those of the putative mothers, 3) individuals with an ambiguous or failed amplification were re-genotyped and 4) all individual genotypes were scored twice, and the scores compared. The program MICROCHECKER version 2.2.0 (van Oosterhout *et al.* 2004) was used to detect any genotyping errors including typing mistakes; extreme stuttering that might obscure heterozygotes, null alleles and large allele dropout.

2.3.5 Genetic diversity analyses

The Excel add-in MS TOOLS (Park 2001) was used to convert all allelic data into appropriate input formats for different data analysis programs. Mean number of alleles per locus, allele frequencies, observed (H_O) and expected (H_E) heterozygosities, polymorphic information content (PIC) and two parentage average exclusion probabilities (P_{E1}: the probability of excluding a parent at a locus or set of loci when lacking information on either of the parents; P_{E2}: the probability of excluding a parent at a locus or set of loci when one parent is known) and null allele frequencies were calculated in CERVUS 2.0 (Marshall *et al.* 1998). The program Doh (Brzustowski 2002) (http://www2.biology.ualberta.ca/jbrzusto/Doh.php) was used to calculate the unbiased estimator of probability of identity P_{ID} (Paetkau *et al.* 1998), the probability of two randomly drawn individuals of having identical genotypes. Deviations from Hardy-Weinberg equilibrium (HWE) within each and over all sampling sites, and tests for linkage disequilibrium (LD) across all pairs of loci were calculated using FSTAT 2.9.3 (Goudet 2002, updated from Goudet 1995) and a Bonferroni correction applied after multiple comparisons (Rice 1989).

2.4 Results and Discussion

2.4.1 Multiplex kit

From the 44 microsatellite loci chosen from the literature, 34 were selected for PCR trials as they had (pre or post re-design) an optimal annealing temperature between 58-65°C and relatively low primer-primer dimerasition. Twenty-one of these primer pairs, 16 with dinucleotide repeats and 5 with tetranucleotide repeats, were further selected because of their good amplification, unambiguous allele scoring and sufficiently different allele size range optimal for multiplex PCR and multimixing. Table 1.1 lists the characteristics and primer sequences for the 21 microsatellite loci chosen, the multiplex PCR sets, multimixing sets, and primer concentrations. In total, 19 of the primer pairs were amplified in five different multiplex PCR reactions: one hexaplex, one pentaplex, one tetraplex and two duplexes. The two other primer pairs

(*OarCP26* and *OarFCB193*) were amplified in separate reactions of 15μL using standard reagents and following the same PCR cycle as for the multiplexed reactions. All 21 microsatellite loci were combined in three multimixing sets, containing 8, 7 and 6 loci, respectively (Table 2.1).

2.4.2 Genotyping and scoring

Mean proportion of individuals typed was 0.987, with 707 individuals with complete genotype for all the 21 microsatellite loci, demonstrating the efficiency of this multiplex kit for large scale genotyping analyses. The five tetranucleotide loci isolated from wapiti (*Cervus elaphus canadensis*): *T26*, *T156*, *T193*, *T501* & *T268* (Jones *et al.* 2002), did not follow a strict 4bp repeat difference between alleles but were a di-tetranucleotide complex indicating an imperfect repetition of the basic motif in red deer. Sequencing of the alleles from these loci would provide more insights into the repeat characteristics of these loci in red deer.

Table 2.1 Characteristics of the multiplex kit

Loading plex	PCR plex	Locus	Origin	Primer sequences	Tm	% GC	Primer concentration (μM)	Size range(bp)	Label
Panel A	Uniplex	*OarCP26*[1]	Sheep	5' – GGCCTAACAGAATTCAGATGATGTTGC– 3'	64.6	44.4	0.4	122-168	VIC
				5' – CCATACTGACGGCTGGTTCC– 3'	64.5	60			
	Uniplex	*OarFCB193*[2]	Sheep	5' – TTCATCTCAGACTGGGATTCAGAAAGG– 3'	64.6	44.4	0.4	102-144	PET
				5' – GCTTGGAAATAACCCTCCTGCATC– 3'	64.5	50			
	PLEX 1	*OarFCB5*[3]	Sheep	5' – AAGTTAATTTTCTGGCTGGAAAACCCCAG– 3'	64.6	41.4	0.2	80-112	6-FAM
				5' – ACCTGACCCTTACTCTCTTCACTC– 3'	64.5	50			
		OarFCB304[2]	Sheep	5' – CCCTAGGAGCTTTCAATAAAGAATCGG	64.6	44.4	0.1	126-146	NED
				5' – CTGCTGTCAACTGGGTCAGG– 3'	64.5	60			
		CelJP38[4]	Red deer	5' – GCTCCAGATTATTCCAGTGTATTGCC– 3'	64.6	46.2	0.1	202-232	VIC
				5' – CTGCACAGAGTCGGACACAAC– 3'	64.5	57.1			
		RT1[7]	Reindeer	5' – CATATGGCTAACTACCTAGCTTGCC– 3'	64.6	48	0.1	252-288	6-FAM
				5' – GAGTCCCAAAGATTTCAGCCCTAC– 3'	64.5	50			
		RT7[7]	Reindeer	5' – CTTTGCCCTGTTCTACTCTTCTTCTC– 3'	64.6	46.2	0.4	213-233	PET
				5' – GCATGGTTTAGGCCCTTGG– 3'	64.5	60			
		TGLA94[6]	Cattle	5' – CATCAAAACAGTGAAGGATGATTGCCAG– 3'	64.6	42.9	0.3	123-155	6-FAM
				5' – CGAATCTCTTCTAGGGATTGAGACTG– 3'	64.5	46.2			
Panel B	PLEX 2	*RT25*[7]	Reindeer	5' – TGCCAAGGAACCAAGATGTC– 3'	60.4	50	0.2	195-207	PET
				5' – CCATTCCAGTATTATTGCCTG– 3'	59.47	42			
		T268[5]	Wapiti	5' – GATGATAACAGCTCAACAGAT– 3'	57.52	38	0.2	207-271	NED
				5' – ATTCCCTTCTCCAGTGTATG– 3'	58.35	45			
		BM757[8]	Cattle	5' – TGGAAACAATGTAAACCTGGG– 3'	59.47	42	0.2	161-217	6-FAM
				5' – TTGAGCCACCAAGGAACCC– 3'	57.89	57			
		RM188/D4S22[9]	Cattle	5' – GCACTATTGGGCTGGTGATT – 3'	60.4	50	0.4	117-141	VIC
				5' – GGTTCACAAAGAGCTGGAC– 3'	60.4	52			
	PLEX 3	*BMC1222*[8]	Cattle	5' – CCAATTTTGCAGATAAGAAAA– 3'	53.62	28.57	1	310-340	6-FAM
				5' – CCTGAGTGTTCCTCCTGAGT– 3'	62.45	55			
		CSSM003[10]	Cattle	5' – GTACCTTAAGGTCAAGGGCTTTCT– 3'	63.7	45.83	0.4	203-233	VIC
				5' – TGGGTCCAATTGAGAATCTTCATG– 3'	61.98	41.66			
Panel C	PLEX 4	*T26*[5]	Wapiti	5' – TGCCATAGTTTTTCCTACCTTC– 3'	59.67	40	0.2	316-268	6-FAM
				5' – GAAGTTCCAATAGACACGCTC– 3'	61.42	47			
		T156[5]	Wapiti	5' – ATGAATACCCAGTCTTGTCTG– 3'	59.47	42	0.3	131-227	6-FAM
				5' – TCTTCCTGACCTGTGTCTTG– 3'	60.4	50			
		T501[5]	Wapiti	5' – CTCCTCATTATTACCCTGTGA– 3'	59.47	42	0.2	172-238	PET
				5' – ACATGCTTTGACCAAGACCC– 3'	60.4	50			
		T193[5]	Wapiti	5' – CTGCTGTTGTCATCATTACCA– 3'	59.47	42	0.4	231-271	NED
				5' – CAGTCCAAGCCTGCTAAATAA– 3'	59.47	42			
		BM888[8]	Cattle	5' – ACTAGGAGGCCATATAGGAGGC– 3'	64.5	57.1	0.1	174-248	PET
				5' – AGCTCAAAACGAGGGACAGGG– 3'	62.5	55			
	PLEX 5	*RT13*[7]	Reindeer	5' – GCCCAGTGTTAGGAAAGAAGA– 3'	61.42	47.62	0.3	176-196	NED
				5' – CATCCCAGAACAGGAGTGAG– 3'	62.45	55			
		CelJP27[4]	Red deer	5' – GCAAATCAGAAATAGACCCACAGAC – 3'	62.9	44	0.1	292-338	VIC
				5' – GATCCCCTCCTTGTGCCAC– 3'	64.5	63.16			

References: [1] Ede et al. (1995); [2] Buchanan & Crawford (1993); [3] Buchanan et al. (1994); [4] Marshall et al. (1998); [5] Jones et al. (2002); [6] Georges & Massey (1992); [7] Wilson et al. (1997); [8] Bishop et al (1994); [9] Barendse et al. (1994); [10] Moore et al. (1994). Primer sequences that do not coincide with published original sequences were specifically re-designed for this study.

When running MICROCHECKER 2.2.0 analysis for each sampling site separate, no extreme stuttering or large allelic dropout was found for any of the loci across all sampling sites. However, there was some indication that some null alleles could be present for loci *RT13* and *T156*. Null allele frequencies obtained in CERVUS 2.0 for these two loci were 0.1620 and 0.1440, respectively. However, detection of null alleles for loci *RT13* and *T156* was not consistent across all sampling sites, with 8 of the 14 sites appearing to have null alleles for both loci and only 5 of the 14 sites when using a reduced data set excluding foetuses genotypes (data not shown). Additionally, no departures from Hardy-Weinberg equilibrium were found for any of the sampling sites indicating that the effect of possible null alleles in loci *RT13* and *T156* might be limited and are not likely to markedly affect population structure analyses in particular in this study where individuals were genotyped for 19 further loci.

2.4.3 Genetic diversity analyses

Individuals presented moderate to high levels of variability at the 21 microsatellite loci, with number of alleles per locus ranging from 7 (*RT25*) to 30 (*BM888*), with 14.43 mean number of alleles per locus (Table 2.2). Mean expected heterozygosity (H_E) and observed heterozygosity (H_O) across all samples were 0.804 and 0.762 respectively, and mean PIC was 0.781. Genetic diversity analyses at the population level revealed similar genetic diversity for all populations (Table 2.3). The measures obtained in this study are comparable and often higher to those found in other population genetic studies of red deer (e.g. Røed 1998; Martínez *et al.* 2002; Kuehn *et al.* 2003; Feulner *et al.* 2004; Kuehn *et al.* 2004; Haanes *et al.* 2005).

The probability of identity was extremely low, $P_{ID} = 2.63 \times 10^{-27}$ (Table 2.2.), indicating a high probability of an individual having an unique genotype and thus making this multiplex kit an ideal protocol for genotyping red deer for forensic analyses such as the identification of illegally killed animals. Parentage exclusion probabilities over all loci were very high $P_{E1} = 0.9999$ and $P_{E2} = 1$ making also this multiplex kit a powerful tool for paternity analysis of red deer.

Table 2.2 Genetic diversity analyses for each of the 21 loci comprising the multiplex kit (n = 871 red deer samples) per locus

Locus	k	H_O	H_E	PIC	P_{E1}	P_{E2}	P_{ID}	Null freq
CP26	17	0.651	0.647	0.625	0.267	0.457	0.142	-0.0062
FCB5	13	0.746	0.764	0.729	0.378	0.555	0.088	0.0114
FCB193	18	0.785	0.823	0.801	0.483	0.655	0.052	0.0227
FCB304	10	0.794	0.824	0.8	0.476	0.65	0.053	0.0178
CelJP38	12	0.813	0.834	0.814	0.504	0.673	0.047	0.0121
RT1	15	0.855	0.872	0.858	0.588	0.742	0.029	0.0094
RT7	9	0.768	0.819	0.795	0.468	0.643	0.056	0.0336
TGLA94	14	0.862	0.878	0.865	0.6	0.751	0.026	0.009
BM888	30	0.8	0.825	0.81	0.511	0.68	0.045	0.016
CelJP27	10	0.833	0.85	0.831	0.529	0.696	0.037	0.0089
RT13	17	0.625	0.868	0.852	0.574	0.731	0.031	**0.162**
T26	14	0.836	0.858	0.843	0.558	0.718	0.033	0.0127
T156	20	0.636	0.851	0.833	0.539	0.703	0.039	**0.144**
T193	17	0.804	0.845	0.829	0.54	0.703	0.039	0.0253
T501	14	0.826	0.844	0.827	0.533	0.697	0.039	0.0079
BM757	15	0.674	0.689	0.673	0.319	0.512	0.112	0.0132
BMC1222	11	0.755	0.759	0.728	0.377	0.558	0.065	0.0006
CSSM003	13	0.702	0.744	0.711	0.356	0.538	0.089	0.0287
RM188	11	0.734	0.737	0.705	0.354	0.535	0.1	0.0003
RT25	7	0.665	0.69	0.627	0.256	0.417	0.159	0.0188
T268	16	0.843	0.855	0.838	0.547	0.709	0.037	0.0074

k = Number of alleles; H_O = observed heterozygosity; H_E = expected heterozygosity; PIC = Polymorphic information content; P_{E1} = probability of parentage exclusion with no parents known; P_{E2} = probability of parentage exclusion with one parent known; P_{ID} = Probability of 2 randomly drawn individuals having identical genotype; Null freq = inferred frequency of null alleles (in bold if they are > 0.05)

Table 2.3 Genetic diversity analyses for each of the 21 loci comprising the multiplex kit (n = 871 red deer samples) per population

Population	Sample size	Males (n)	Females (n)	Foetuses (n)	Mean k	A	H_E	H_O	F_{IS}
K	76	29	29	18	10.47	8.30	0.80	0.76	0.058
SA	74	30	30	12	10	8.21	0.79	0.79	0.017
CON	49	19	30	-	9.71	8.35	0.79	0.76	0.049
AG	28	10	18	-	8.67	8.57	0.79	0.76	0.050
MA	56	30	26	-	10.14	8.45	0.79	0.78	0.017
GC	86	30	28	28	9.62	8.11	0.77	0.74	0.048
GCR	33	21	11	-	8.57	8.34	0.77	0.77	0.015
GK	78	52	15	11	9.19	7.85	0.77	0.75	0.030
GST	34	26	8	-	8.43	7.72	0.77	0.79	-0.003
CO	30	30	-	-	8.52	8.54	0.77	0.73	0.063
AR	83	30	27	26	10.10	8.62	0.79	0.78	0.025
BA	66	30	18	18	9.43	8.33	0.79	0.77	0.034
CL	90	30	30	30	10.05	8.07	0.77	0.75	0.042
FL	86	29	28	29	10.24	7.87	0.78	0.77	0.027

Mean k = Mean number of alleles/locus; A = Allelic richness (based on min. sample 24 individuals); H_O = observed heterozygosity; H_E = expected heterozygosity; F_{IS} = heterozygote deficiency

The multiplex protocol described here has proved very useful for high throughput microsatellite genotyping of red deer. This multiplex kit is an efficient and reliable tool for not only evolutionary and ecological studies of populations of red deer but also for more applied studies on conservation or management of red deer populations, and for forensic detection of illegal trafficking of trophies.

2.5 Acknowledgements

Stalkers and deer manager from all the collaborative estates are greatly thanked for the collection of samples. W.P. Goodall-Copestake is thanked for his advice on developing the multiplex kit. Comments from F.J. Pérez-Barbería, J.M. Pemberton, C.D. Jiggins and I.J. Gordon greatly improved this manuscript. A. Gillies and J. Lovell are thanked for their sequencer services at IEB. The Macaulay Development Trust funded this study and supported S.P.-E. with a studentship.

2.6 References

ArKdb database. http://www.thearkdb.org

Barendse W, Armitage D, Kossarek L, *et al.* (1994) A genetic linkage map of the bovine genome. *Nature Genetics* **6**, 2.

Bishop MD, Kappes SM, Keele JW, *et al.* (1994) A genetic linkage map for cattle. *Genetics* **136**, 619-639.

Bonnet A, Thévenon S, Maudet F, Maillard JC (2002) Efficiency of semi-automated fluorescent multiplex PCRs with 11 microsatellite markers for genetic studies of deer populations. *Animal Genetics* **33**, 343-350.

Buchanan FC, Crawford AM (1993) Ovine microsatellite at the OarFCB11, OarFCB128, OarFCB193, OarFCB266 and OarFCB304. *Animal Genetics* **24**, 145.

Buchanan FC, Galoway SM, Crawford AM (1994) Ovine microsatellites at the OarFCB5, OarFCB19, OarFCB20, OarFCB48, OarFCB129, and OarFCB226 loci. *Animal Genetics* **25**, 60.

Brzustowski J (2002) Doh. http://www2.biology.ualberta.ca/jbrzusto/Doh.php

Chamberlain JS, Gibbs RA, Ranier JE, Nguyen PN, Caskey CT (1988) Deletion screening of the Duchenne muscular distrophy locus via multiplex DNA amplification. *Nucleic Acids Research* **16**, 11141-11156.

Ciofi C, Funk SM, Coote T, *et al.* (1998) Genotyping with microsatellite markers. In: *Molecular Tools for Screening Biodiversity* (eds. Karp A, Isaac PG, Ingram DS), pp. 195-205. Chapman & Hall, London.

Clutton-Brock TH, Guinness FE, Albon SD (1982) *Behaviour and Ecology of Two Sexes* University of Chicago Press, Chicago.

Cryer NC, Butler DR, Wilkinson MJ (2005) High throughput, high resolution selection of polymorphic microsatellite loci for multiplex analysis. *Plant Methods* **1**.

Ede AJ, Pierson CA, Crawford AM (1995) Ovine microsatellites at the OarCP9, OarCP16, OarCP20, OarCP21, OarCP23 and OarCP26. *Animal Genetics* **25**, 129-130.

Engel SR, Linn RA, Taylor JF, Davis SK (1996) Conservation of microsatellite loci across species of artiodactyls: implications for population studies. *Journal of Mammalogy* **778**, 504-551.

Feulner PGD, Bielfeldt W, Zachos FE (2004) Mitochondrial DNA and microsatellite analyses of the genetic status of the presumed subspecies *Cervus elaphus montanus* (Carpathian red deer). *Heredity* **93**, 299-306.

Fickel J, Reinsch A (2000) Microsatellite markers for the European roe deer (*Capreolus capreolus*). *Molecular Ecology* **9**, 993-1011.

Galan M, Cosson JF, Aulagnier S, *et al.* (2003) Cross amplification tests of ungulate primers in roe deer (*Capreolus capreolus*) to develop a multiplex panel of 12 microsatellite loci. *Molecular Ecology Notes* **3**, 142-146.

Geist V (1998) *Deer of the world: their evolution, behaviour, and ecology* Stackpole Books, Mechanicsburg, PA.

Georges MM, J.M. (1992) (1992) Polymorphic DNA markers in Bovidae.

Goodman SJ, Barton NH, Swanson GM, Abernethy K, Pemberton JM (1999) Introgression through rare hybridization: A genetic study of a hybrid zone between red and sika deer (Genus *Cervus*) in Argyll, Scotland. *Genetics* **152**, 355-371.

Goudet J (1995) FSTAT version 1.2: a computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Haanes H, Rosef O, Veiberg V, Røed KH (2005) Microsatellites with variation and heredity applicable to genetic studies of Norwegian red deer (*Cervus elaphus atlanticus*). *Animal Genetics* **36**, 454-455.

Hammond RL, Saccheri IJ, Ciofi *et al.* (1998) Isolation of microsatellite markers in animals. In: *Molecular Tools for Screening Biodiversity: Plants and Animals* (eds. Karp A, Isaac PG, Ingram DS). Chapman & Hall, London.

Hartl G, Zachos F, Nadlinger K (2003) Genetic diversity in European red deer (*Cervus elaphus* L.): anthropogenic influences on natural populations. *Comptes Rendus Biologies* **326**, 37-42.

Hartl GB, Willing R, Lang G, Klein F, Köller J (1990) Genetic variability and differentiation in red deer (*Cervus elaphus* L.) of Central Europe. *Genetics Selection Evolution* **22**, 289-306.

Henegariu O, Heerema NA, Dlouy SR, Vance GH, Vogt PH (1997) Multiplex PCR: Critical parameters and step-by-step protocol. *Biotechniques* **23**, 504-511.

Jones KC, Levine KF, Banks JD (2002) Characterization of 11 polymorphic tetranucleotide microsatellites for forensic applications in California elk (*Cervus elaphus canadensis*). *Molecular Ecology Notes* **2**, 425-427.

Keyghobadi N, Roland J, Strobeck C (2005) Genetic differentiation and gene flow among populations of the alpine butterfly, *Parnassius smintheus*, vary with landscape connectivity. *Molecular Ecology* **14**, 1897-1909.

Kuehn R, Anastassiadis C, Pirchner F (1996) Transfer of bovine microsatellites to te cervine (*Cervus elaphus*). *Animal Genetics* **27**, 199-201.

Kuehn R, Haller H, Schroeder W, Rottman O (2004) Genetic roots of the Red deer (*Cervus elaphus*) in Eastern Switzerland. *Journal of Heredity* **95**, 136-143.

Kuehn R, Schroeder W, Pirchner F, Rottmann O (2003) Genetic diversity, gene flow and drift in Bavarian red deer populations (*Cervus elaphus*). *Conservation Genetics* **4**, 157-166.

Lowe VPW, Gardiner AS (1974) A re-examination of the subspecies of Red deer (*Cervus elaphus*) with particular reference to the stocks of Britain. *Journal of Zoology (London)* **174**, 185-201.

Luikart G, England P (1999a) Statistical analysis of microsatellite DNA data. *Trends in Ecology and Evolution* **14**, 253-256.

Luikart G, Biju-Duval MP, Ertugrul, *et al.* (1999b) Power of 22 microsatellite markers in fluorescent multiplexes for parentage in goats (*Capra hircus*). *Animal Genetics* **30**, 432-438.

Maddox JF, Davies PK, Crawford AM, Hulme DJ, *et al.* (2001) An enhanced linkage map of the sheep genome comprising more than 1000 loci. *Genome Research* **11**, 1275-1289.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639-655.

Martínez JG, Carranza J, Fernández-García JL, Sánchez-Prieto CB (2002) Genetic variation of red deer populations under hunting exploitation in southwestern Spain. *Journal of Wildlife Management* **66**, 1273-1282.

Moore SS, Sargeant LL, King TJ, *et al.* (1991) The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics* **10**, 654-660.

Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution* **9**, 373-375.

Niemczewski C, Rutkowski J, Zurkowski M (2002) Preliminary investigations on the polymorphism of microsatellite markers in Mazurian red deer (*Cervus elaphus*). *Animal Science Papers and Reports* **20**, 169-174.

Oligotoolkit, Operon. http://www.operon.com/oligos/toolkit.php

Paetkau D, Waits LP, Clarkson PL, *et al.* (1998) Variation in genetic diversity across the range of North American brown bears. *Conservation Biology* **12**, 418-429.

Park SDE (2001) *Trypanotolerance in West African cattle and the population genetic effects of selection*, University of Dublin.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Røed KH (1998) Microsatellite variation in Scandinavian Cervidae using primers derived from Bovidae. *Hereditas* **129**, 19-25.

Røed KH, Midrhjell L (1998) Microsatellites in reindeer, *Rangifer tarandus*, and their use in other cervids. *Molecular Ecology* **7**, 1771-1788.

Schlötterer C (2002) Microsatellites. In: *Molecular Genetic Analysis of Populations* (ed. HOELZEL AR), pp. 237-261. Oxford Press University, New York.

Slate J, Coltman DW, Goodman SJ, *et al.* (1998) Bovine microsatellite loci are highly conserved in red deer (*Cervus elaphus*), sika deer (*Cervus nippon*) and Soay sheep (*Ovis aries*). *Animal Genetics* **29**, 307-315.

Slate J, Van Stijn TC, Anderson RM, *et al.* (2002) Genetic linkage map and the evolution of ruminant genomes. *Genetics* **160**, 1587-1597.

Ströhlein H, Herzog S, Hecht W, Herzog A (1993) Biochemical genetic description of German and Swiss populations of red deer *Cervus elaphus*. *Acta Theriologica Suppl.* **2**, 153-161.

Thomas MG, Bradman N, Flinn HM (1999) High throughput analysis of 10 microsatellite and 11 diallelic polymorphisms on the human Y-chromosome. *Human Genetics* **105**, 577-581.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

Vial L, Maudet C, Luikart G (2003) Thirty-four polymorphic microsatellites for European roe deer. *Molecular Ecology Notes* **3**, 523-527.

Wilson GA, Strobeck C, Wu L, Coffin J (1997) Characterization of microsatellite loci in caribou Rangifer tarandus, and their use in other artiodactyls. *Molecular Ecology Notes* **6**, 697-699.

Ziegle JS, Su Y, Corcoran KP, *et al.* (1992) Application of automated sizing technology for genotyping microsatellite loci. *Genomics* **14**, 1026-1031.

Chapter 3

Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*)

S. Pérez-Espona,[*§] F.J. Pérez-Barbería, [§] J. McLeod, [§] C.D. Jiggins, [*] I.J. Gordon, † J. M. Pemberton[*]

*Institute of Evolutionary Biology, School of Biological Sciences,The University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

§ The Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK

† Sustainable Ecosystems, CSIRO – Davies Laboratory, PMB PO Aitkenvale, Qld 4814, Australia

Corresponding author:

Sílvia Pérez-Espona,

Institute of Evolutionary Biology, School of Biological Sciences, The University of Edinburgh

The King's Buildings, West Mains Road, Edinburgh EH9 3J, UK

Fax: ++ 44 (0)131 650 6564

Tel: ++44 (0)131 651 3612

E-mail: s.perez-espona@sms.ed.ac.uk

Author's contribution:

SPE organised the sampling, carried out the laboratory work, performed the statistical analysis and wrote the manuscript. JM constructed GIS least-cost distance matrices and landscape maps. The project was set up by FJPB and IJG. The whole project was supervised by FJPB, JMP, CDJ and IJG. All authors critically read and improved drafts of this manuscript.

3.1 Abstract

Landscape features have been shown to strongly influence dispersal and consequently the genetic population structure of organisms. Studies quantifying the effect of landscape features on gene flow of large mammals with high dispersal capabilities are rare and have mainly focused at large geographical scales. In this study we assessed the influence of several natural and man-made landscape features on red deer gene flow in the Scottish Highlands by analysing 695 individuals for 21 microsatellite markers. Despite the relatively small scale of the study area, significant population structure was found using F-statistics (F_{ST} = 0.019) and the program STRUCTURE (K= 4), with major differentiation found between populations sampled at either side of the Great Glen. Following a landscape genetics approach, GIS techniques were used to construct least-cost distance matrices using a range of cost values to cross a particular landscape feature. Landscape features were shown to significantly affect red deer gene flow as they explained a greater proportion of the genetic variation than the geographical distance between populations. Sea lochs were found to be the strongest red deer gene flow barriers in our study area, followed by roads, mountain slopes and forests. Inland lochs and rivers were identified as red deer gene flow corridors. Additionally, we demonstrate the impact of choosing arbitrary cost values to construct least-cost distance matrices and how the interaction between several landscape features needs to be taken account to understand the effect of landscape on population differentiation.

Keywords:

Cervus elaphus, dispersal, landscape genetics, least-cost distance, GIS, microsatellites, population structure, red deer

3.2 Introduction

Effective dispersal (gene flow) is one of the major evolutionary processes influencing population genetic structure (Clobert *et al.* 2001). In nature, populations are frequently subdivided into distinct subpopulations connected by different degrees of gene flow. The use of population genetic studies to understand how dispersal affects genetic variation within and between populations through gene flow can not only provide new insights about the evolutionary biology and ecology of species but it can also provide valuable information for conservation and management policies (Crandall *et al.* 2000; Moritz 2002). However, gene flow between populations is a complex process that is influenced by many intrinsic factors such as the innate dispersal ability and breeding system of an individual, and also by extrinsic factors such as landscape features or other environmental factors (Lowe *et al.* 2004).

The influence of landscape features on dispersal and ultimately on population genetic structure has long been recognised (Fisher & Ford 1947). However, only recently, with the emergence of the new discipline of landscape genetics, which combines landscape ecology and population genetics, has the effect of landscape features on genetic population structure been assessed in a quantitative way (Manel *et al.* 2003). Although putative gene flow barriers can be revealed by associating the presence of certain landscape features with observed genetic discontinuities, it is often difficult to separate which particular features are the most likely to be affecting dispersal due to the strong correlations between individual landscape features in nature. The incorporation of geographical information systems (GIS) and geostatistics into landscape genetic studies is providing useful tools to visualise and quantify the effect of individual and/or combinations of landscape features on the observed population structure in a variety of organisms (e.g. Piertney *et al.* 1998; Funk *et al.* 2005; Spear *et al.* 2005; Vignieri 2005).

For medium to large size mammals, studies on the effect of landscape features on population genetic structure have mainly been assessed on carnivores and have

focused on very large geographical areas (e.g. Geffen *et al.* 2004; McRae *et al.* 2005; Sacks *et al.* 2005). For large non-carnivorous mammals with much narrower ranges, landscape genetics studies are rare. In ungulates, landscape features acting as gene flow barriers have been suggested or indirectly inferred from observed spatial population genetic structure (e.g. red deer, Hartl *et al.* 2005; roe deer, *Capreolus capreolus*, Coulon *et al.* 2006). However, the influence of landscape features on population structure has only rarely been assessed using a direct quantitative approach (e.g. Coulon *et al.* 2004; Epps *et al.* 2005). Using a forest numeric model Coulon *et al.* (2004) found that roe deer female gene flow was significantly linked to forests in their study area in southern France. Epps *et al.* (2005), incorporating the effects of human-made barriers in isolation by distance regressions found that highways, canals and developed areas less than forty years old were causing a rapid decrease of gene flow between desert bighorn sheep (*Ovis canadensis nelsoni*) populations in southern California.

Red deer (*Cervus elaphus*) is the largest terrestrial mammal in Britain and it is widely distributed across Scotland, with the highest densities being found in the Central and Eastern Highlands (Clutton-Brock & Albon 1989). Red deer are considered to be highly mobile animals, capable of not only dispersing large distances but also crossing many kinds of terrain, including swimming across large water bodies. Nonetheless, it is expected that the crossing of some sorts of terrain might require a higher energetic cost than others, which could prevent individuals from achieving maximum dispersal distances and thus result in lower dispersal rates between some geographical areas (Sugg *et al.* 1996; Dobson 1998). In fact, red deer mark-recapture data from different calf tagging programs (with tags recovered from dead or shot individuals), carried out at different periods of times between 1967-2001 on the Scottish mainland, have shown that although some individuals might disperse long distances (> 50 km), mean overall dispersal distances are commonly short, 3.3-7.4 km for males and 1.9-3.5 km for females (Daniels & McClean 2003). In addition, a study in the Cairngorms area following 18 adult males and 4 females carrying Global Positioning System (GPS) collars showed that, on average, males dispersed around 10 km from their normal over-wintering area (3-21 km) during the breeding

season whereas hinds moved on averaged less than 3 km (Sibbald *et al.* unpublished data).

Due to the fact that red deer has been heavily influenced by humans during the last 300 years, either by hunting or introduction of individuals from foreign stocks (Lowe & Gardiner 1974; Clutton-Brock *et al.* 1989; Hartl *et al.* 2003), we tested if any of our studied populations has recently undergone a genetic bottleneck. Despite the capability of red deer of moving through different types of terrain, we hypothesise that the crossing of some landscape features might imply a much higher energetic cost and thus we would expect lower levels of gene flow between areas separated by certain landscape features. To test this hypothesis we took a landscape genetics approach and used GIS techniques to assess the influence of several natural and man-made landscape features (sea lochs, inland lochs, rivers, railways, forests, mountain slopes and roads) on red deer gene flow. The effect of choosing arbitrary cost values to construct least-cost distance matrices and the interaction of several landscape features on red deer gene flow are also discussed. Finally, the outcome of this study will also provide useful information for current red deer management in mainland Scotland. The current red deer range in Scotland is divided into several Deer Management Groups, areas which are thought to contain distinct populations or herds of red deer and are generally delimited by considering natural or man-made geographical features.

3.3 Materials and Methods

3.3.1 Study area and sampling

The study area comprises a series of adjacent or closely located estates distributed across an east-west transect in the Scottish Highlands (Fig. 2.1), with a maximum distance of c.100 km between midpoints of the most distant estates (FL and K). In total 14 estates were sampled, four west of the Great Glen (K, SA, CON, AG) and ten east of the Great Glen (FL, CL, BA, AR, CO, MA, GC, GCR, GK, GST) (Full names are given in the legend of Fig.2.1). The Great Glen is a 97 km long valley that runs in a

SW-NE direction, from Fort William at the head of the sea loch Loch Linnhe to Inverness on the Moray Firth (Fig. 2.1). It is located on the Great Glen Fault, a strike-slip fault that splits the Scottish Highlands into two parts. More than half of the length of the Great Glen is occupied by water, with the freshwater lochs (lakes) of Loch Lochy, Loch Oich and Loch Ness and a navigable waterway, the Caledonian Canal, linking the Atlantic at Loch Linnhe with the North Sea at Inverness, via the three lochs (Whittow 1992; Fig. 2.1).

Red deer samples consisted of either an ear tip or a jaw from wild individuals legally shot (culled) in open hill habitats during the 2003-2004 hunting season. In total 699 individuals were collected (400 males and 299 females). Additional information such as death date, death location, estimated age and carcass weight was provided for each sample.

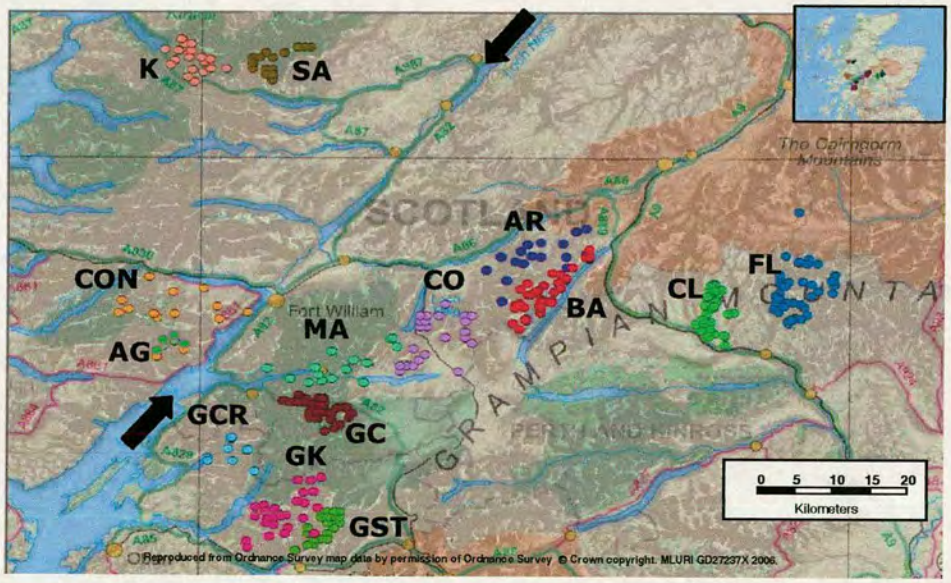


Fig. 3.1 Map of the sampling area. Dots in the map do not correspond to the total of individuals sampled but to the culling areas (several individuals were sampled in the same location). Black arrows indicate location of the Great Glen. **FL** = Forest Lodge (n = 57), **CL**= Clunes (n = 60), **BA**= Ben Alder (n = 48), **AR**= Ardverikie (n = 57), **CO=** Corrour (n = 30), **MA**= Mamore (n = 56), **GC**= Glencoe (n = 58), **GCR**= Glencreran (n = 33), **GK**= Glenkinglass (n = 67), **GST**= Glenstrae (n = 34), **CON**= Conaglen (n = 49), **AG**= Ardgour (n = 28), **SA**= South Glen Affric, (n = 60), **K**= Kintail (n = 58).

3.3.2 DNA laboratory procedures

Tissue samples were stored either in a -20ºC freezer or in tubes containing 100% ethanol. Genomic DNA was extracted from ear or jaw muscle using the DNAace Spin Tissue Mini Kit (Bioline) or with the DNEasy Tissue Kit™ (QIAGEN), following the manufacturer's instructions. A 21-microsatellite loci high-throughput protocol developed for PCR and sequencer load multiplexing was used to genotype all the individuals (Chapter 2). Multiplex PCR products were run on an ABI 3730 capillary sequencer (Applied Biosystems) together with the internal size standard GeneScan 500 LIZ (Applied Biosystems). Fragment analysis was conducted using the software GeneMapper™ v. 3.0 (Applied Biosystems).

3.3.3 Data analysis

All 699 samples were genotyped; four redundant genotypes (one pair of samples with completely identical genotype for all loci and three pairs of samples with all but one locus matching) were excluded from all the analyses leaving a final data set consisting of 695 individuals. The program MICROCHECKER v.2.2.0 (Van Oosterhout *et al.* 2004) was used to detect any genotyping errors (including typing mistakes), extreme stuttering that might have obscured heterozygotes, null alleles and large allele dropout.

3.3.3.1 Genetic diversity analyses

The Excel add-in MS TOOLS (Park 2001) was used to convert all allelic data into several different input formats for different data analysis programs. Genetic diversity measures such as mean number of alleles per locus, allele frequencies, observed (H_O), expected (H_E) heterozygosities, polymorphism information content (PIC) and null allele frequencies were calculated using CERVUS v.2.0 (Marshall *et al.* 1998).

Allele richness, the number of alleles corrected for the minimum sample size (24 individuals), was estimated using FSTAT v.2.9.3 (Goudet 2001 updated from Goudet 1995). Deviations from Hardy-Weinberg equilibrium (HWE) within each and over all sampling sites, and tests for linkage disequilibrium (LD) across all pairs of loci were also conducted using FSTAT, with a Bonferroni correction applied for multiple comparisons ($\alpha = 0.05$, Rice 1989).

The program BOTTLENECK v.1.2 (Cornuet & Luikart 1996; Piry *et al.* 1999) was used to test if any of the populations had undergone a reduction in their effective population size. Deviation of expected heterozygosity under a mutation-drift equilibrium model was calculated using a two-phase mutation model (TPM), an intermediate model between the infinite allele model (IAM) and the stepwise mutation model (SMM) as it seems to perform better for microsatellites (Di Rienzo *et al.* 1994; Spencer *et al.* 2000). Significance values were obtained using a Wilcoxon signed rank test with 1000 iterations (Luikart *et al.* 1998). Additionally, the qualitative descriptor "mode-shift" was run to discriminate stable populations from populations that could have suffered a bottleneck (Luikart *et al.* 1998). Populations that have remained stable through time should show an L-shaped allele frequency distribution as expected under the mutation-drift equilibrium model whereas a bottlenecked population would show a mode-shift.

3.3.3.2 Population differentiation analyses

Population differentiation was assessed using two approaches, first by using traditional F-statistics (Wright 1978) comparing the observed averaged allele distributions among *a priori* defined populations. Although robust population structure analysis can be obtained with F-statistics (Neigel 1997), the *a priori* subdivision of populations can be very subjective and may therefore have a significant effect on the estimates of population structure (Balloux & Lugon-Moulin 2002). The second approach taken was by using an individual-based Bayesian clustering method implemented in the program STRUCTURE v.2.1 (Pritchard *et al.* 2000) in which the number of potential genetic clusters (K) are inferred without

defining populations *a priori*. F-statistics across the study area and between pairs of sampling sites were calculated using the estimator θ of Wright's F_{ST} (Weir & Cockerham 1984) using the program GENEPOP v.3.4 (Raymond & Rousset 1995). Significance values over all loci were obtained using a Fisher's exact test (Ryman & Jorde 2001) with 10,000 dememorisation steps, 100 batches and 500 iterations per batch in the Markov chain method of Guo & Thompson (1992). For the Bayesian clustering method analysis in STRUCTURE, five independent runs for K= 1-14 (total number of estates sampled) were conducted to ensure consistency of the inferred K and convergence of parameters, as recommended by the authors (Pritchard *et al.* 2000). K was calculated using a burn-in of 30,000 replications and 10^6 MCMC steps assuming a model of admixture and correlated frequencies among populations.

3.3.3.3 Effect of geographical distance and landscape features on red deer population structure

Isolation by distance

To investigate if the differentiation between sampling sites followed an isolation by distance model (Wright 1943; Slatkin 1993), Mantel tests (Mantel 1967) between pair-wise genetic distances expressed as $F_{ST}/(1-F_{ST})$ against the natural log-transformed geographical distance (Rousset 1997) were performed in GENEPOP. Statistical significance was estimated after 10,000 permutation tests. Isolation by distance was tested at different scales: the whole study area, between sampling sites located east of the Great Glen, and between sampling sites located west of the Great Glen.

Detection of gene flow barriers for red deer gene flow

Genetic discontinuities between the studied populations that could be attributed to the presence of gene flow barriers were investigated by calculating in FSTAT the correlation of geographical and genetic distance only between estates located at

either side of the Great Glen (as described above). The program BARRIER v.2.2 (Manni *et al.* 2004) was also used to visualise the location and strength of any genetic discontinuities. This program uses a geometric based computation method to detect areas where abrupt changes in genetic distances occur. To remove any effect of isolation by distance between sampling sites, the genetic distance matrix consisted of the residuals from a regression analysis between geographical distance and the multilocus pair-wise F_{ST} matrix. As bootstrap matrices were not available, the robustness of the barrier was alternatively calculated by plotting the thickness (robustness) of the barrier proportionally to the inverse ratio between the maximum F_{ST} distance and the F_{ST} values associated with crossed edges as recommended by the authors (Manni *et al.* 2004).

Effect of landscape features on red deer population structure

To quantify the effect of several natural and man-made landscape features on red deer gene flow we took a landscape genetics approach using GIS techniques. Mid location points for each of the 14 populations were calculated using the geographical coordinates for each individual sampling location and then plotted on a digitised map using ARCMAP v.9.1 (ESRI). The following landscape features were selected to assess their influence as corridors or barriers to red deer gene flow: sea lochs, inland lochs, rivers, railways, roads, forests and mountain slopes (Fig. 3.2). Digitised maps for each of these landscape features were obtained from the Digital Elevation Model (slopes), the Ordnance Survey Strategi (roads, rivers and coastlines) and the Land Cover of Scotland from 1988 (inland lochs and forests). Railways were manually digitised by tracing the railway lines on a 1:1,000,000 scale digital OS maps (Miniscale). Least-cost distance matrices were calculated for each landscape feature by considering each map as a grid and assigning to each cell (50 x 50 m) on the grid a cost (the cost for an individual to cross that cell). All cells were assigned a cost = 1, except for those containing the landscape feature of interest for which a cost <1 or >1 was assigned if the landscape feature was considered a dispersal corridor or a barrier, respectively. Then, all possible directions to move between populations were

calculated using the ArcGIS CostDistance function and the least costly path recorded. Due to the arbitrariness of cost values, the lack of information on which landscape features act as corridors or barriers for red deer dispersal and the relative energetic cost for red deer to cross a certain landscape feature, we assessed a range of cost values (0.1, 0.32, 1, 3.2, 10, 32, 100, 320, 1000, 3200) so that the logarithm of the cost values increased in step sizes. Least cost-distance matrices constructed with cost = 1 represented the Euclidean (straight-line) distance between populations. To decide the most appropriate cost value for each landscape feature we conducted a regression analysis using the population pair-wise F_{ST} distance matrix as dependant variable and each of the different least-cost distance matrices as explanatory variable. Simple Mantel tests were run in FSTAT and significance values obtained after 10,000 randomisations.

To assess the relative effect of each landscape feature in red deer gene flow we conducted partial Mantel tests between the pair-wise F_{ST} matrix and each of the landscape features after correcting for the amount of variation already explained by distance alone (as all landscape feature matrices contained the effect of distance). Partial Mantel tests were also calculated using FSTAT and significance values were obtained after 10,000 randomisations. As the genetic differentiation between populations might be affected by the presence of several landscape features (corridors and barriers) in a particular area, it is important to assess the interaction between landscape features. To examine the interaction between landscape features we conducted additional partial Mantel tests in FSTAT as described above (still correcting by distance) but including two landscape least-cost distance matrices. All possible combinations of pairs of landscape matrices (including order permutations) were tested to confirm the interaction. Partial Mantel tests including more than two landscape variables exponentially increased the number of tests to conduct and are not shown here as the interaction of landscape features can already be appreciated when combining only two landscape variables.

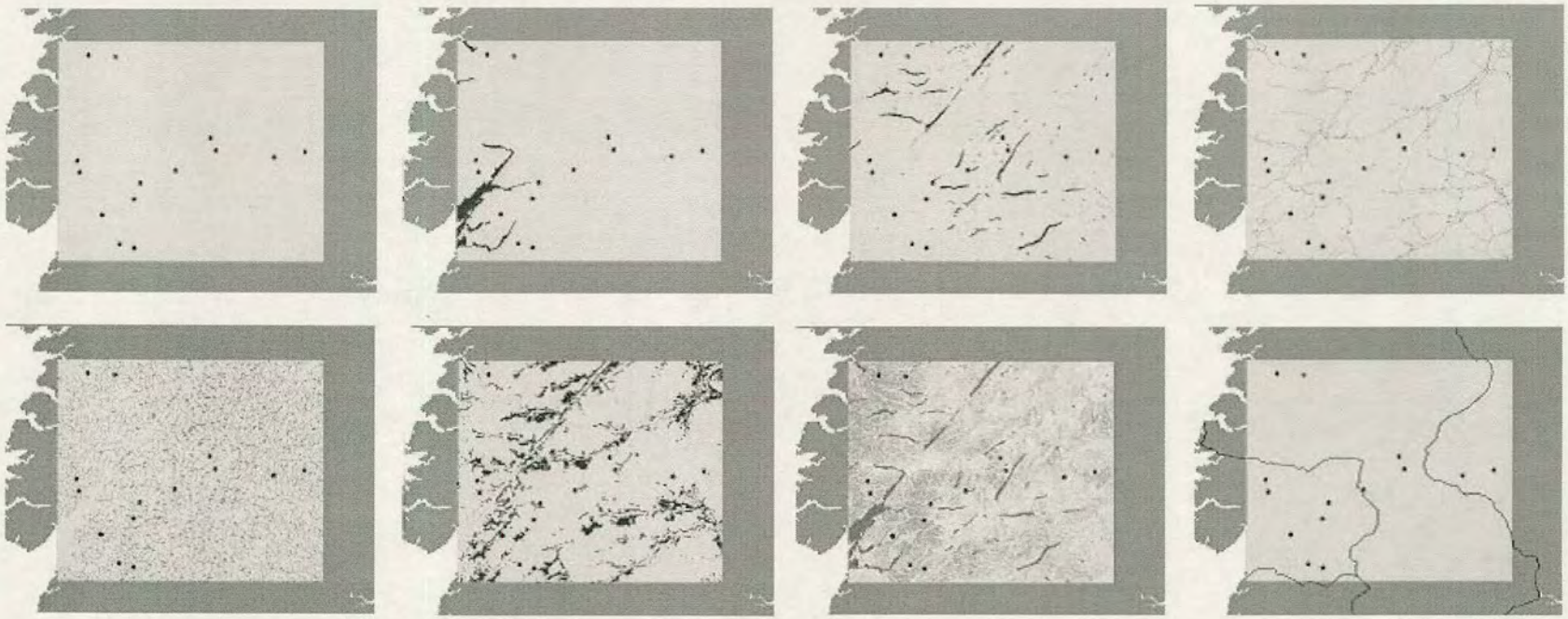


Fig. 3.2 Digitised maps for several landscape features that might affect red deer gene flow in the Scottish Highlands. From left to right and top to bottom: Euclidean distance, sea lochs, inland lochs, roads, rivers, forests, slope and railways. All maps are reproduced from Ordnance Survey map data by permission of Ordnance Survey © Crown copyright. MLURI GD27237X 2006.

3.4 Results

3.4.1 Genetic diversity

Results from the genetic diversity analyses are summarised in Table 3.1. The individuals sampled in this study showed moderate to high levels of variability at the 21 microsatellite loci. The number of alleles per locus ranged from 7 (*RT25*) to 30 (*BM888*), with a mean of 14.29 alleles per locus. Polymorphism Information Content (PIC) ranged from 0.625 (*CP26*) to 0.861 (*TGLA94*), with a mean of 0.781. Mean expected heterozygosity (H_E) and observed heterozygosity (H_O) across the study area were 0.804 and 0.763, respectively. Overall allelic richness ranged from 4.513 (*RT25*) to 12.557 (*BM888*). Possible null alleles were detected for loci *RT13* and *T156*. However, this was not consistent across all sampling sites. Furthermore, no departures from Hardy-Weinberg equilibrium were detected for any of the sampling sites indicating that the effect of possible null alleles in *RT13* and *T156* might be limited and thus we included them in our data set (see also Chapter 2). No linkage disequilibrium was detected for any pair of loci. Similar genetic diversity values for all sampling areas were obtained when conducting the analyses at the population level (Table 3.2)

Bottleneck analysis did not provide evidence for any the populations having experienced a reduction in effective population size as none of the Wilcoxon rank tests gave significant deviation from mutation-drift equilibrium and all populations showed a normal L-shaped distribution of allele frequencies (data not shown).

Table 3.1 Genetic diversity analyses for each of the 21 loci comprising the multiplex kit (n = 695 red deer samples) per locus

Locus	K	H_O	H_E	PIC	Null freq
CP26	17	0.651	0.646	0.625	-0.0004
FCB5	13	0.741	0.761	0.725	0.0125
FCB193	18	0.779	0.825	0.803	0.0281
FCB304	9	0.807	0.828	0.804	0.0116
CelJP38	12	0.813	0.803	0.814	0.0190
RT1	15	0.855	0.865	0.856	0.0021
RT7	9	0.768	0.771	0.793	0.0314
TGLA94	14	0.862	0.856	0.861	0.0109
BM888	30	0.8	0.809	0.808	0.0066
CelJP27	10	0.833	0.838	0.832	0.0064
RT13	17	0.610	0.867	0.852	**0.1735**
T26	14	0.851	0.861	0.845	0.0045
T156	20	0.636	0.623	0.832	**0.1539**
T193	17	0.804	0.805	0.833	0.0266
T501	14	0.826	0.834	0.829	0.0048
BM757	15	0.674	0.674	0.671	0.0122
BMC1222	11	0.755	0.764	0.733	-0.0008
CSSM003	12	0.693	0.740	0.706	0.0326
RM188	11	0.739	0.739	0.707	-0.0023
RT25	7	0.663	0.693	0.631	0.0226
T268	16	0.852	0.857	0.840	0.0029

k = number of alleles/locus; PIC = polymorphism information content; H_O = observed heterozygosity; H_E = expected heterozygosity; Null freq = inferred null allele frequency

Table 3.2 Genetic diversity analyses for each of the 21 loci comprising the multiplex kit (n = 695 red deer samples) per population

Population	Sample size	Males (n)	Females (n)	Mean number alleles/locus	Allelic richness	H_E	H_O	F_{IS}
K	58	29	29	10.33	8.63	0.79	0.76	0.068
SA	60	30	30	9.86	8.41	0.80	0.79	0.021
CON	49	19	30	9.71	8.57	0.79	0.76	0.049
AG	28	10	18	8.67	8.45	0.79	0.76	0.050
MA	56	30	26	10.14	8.62	0.79	0.78	0.017
GC	58	30	28	9.48	8.05	0.77	0.75	0.040
GCR	33	21	11	8.57	8.07	0.77	0.77	0.015
GK	67	52	15	9	7.72	0.77	0.74	0.036
GST	34	26	8	8.43	7.87	0.77	0.79	-0.003
CO	30	30	-	8.52	8.11	0.77	0.73	0.063
AR	57	30	27	9.86	8.50	0.79	0.77	0.027
BA	48	30	18	9.19	8.18	0.79	0.77	0.029
CL	60	30	30	9.9	8.48	0.77	0.76	0.034
FL	57	29	28	9.57	8.29	0.78	0.76	0.033

A = Allelic richness (based on min. sample 24 individuals); H_O = observed heterozygosity; H_E = expected heterozygosity; F_{IS} = heterozygote deficiency

3.4.2 Population differentiation analyses

Population differentiation across the whole study area was low ($F_{ST} = 0.019$) but highly significant (99 % confidence limits = 0.015 - 0.022). Using Fisher exact tests and after applying a Bonferroni correction, the number of significant pair-wise comparisons was 81 out of 91, with the following pair of adjacent or closely located sampling sites not being significantly differentiated from each other: K-SA, CON-AG, GC-GK, CO-MA, GST-CO, MA-AR, AR-CO, BA-CO, BA-AR, and CL-FL. Pair-wise F_{ST} values ranged from 0 (CON-AG, AR-BA) to a maximum value of 0.0434 between GCR-AG (Table 3.3).

Table 3.3 Genetic and geographical distances between the 14 Scottish Highland red deer populations sampled. Genetic distance is represented by pair-wise $F_{ST} / (1-F_{ST})$ (lower diagonal) and geographical distance in km (upper diagonal)

Population	K	SA	CON	AG	MA	GC	GCR	GK	GST	CO	AR	BA	CL	FL
K	-	11.6	46.5	51.4	60	65.7	82.4	82.4	86.4	62.5	63.4	68.4	91.1	101.9
SA	0.002	-	40.8	51	58.2	58.5	64.6	74.7	81.2	52.4	50.2	54.7	82	89.9
CON	0.006	0.007	-	9.3	26.3	26.3	27.4	41.3	48.2	45.2	58	55.1	88.3	98.8
AG	0.010	0.009	0	-	25.5	25.6	22.9	37	45.5	49.8	64.1	60.7	93.5	104
MA	0.023	0.023	0.020	0.025	-	18.4	27.7	34.2	38.5	16.9	32.2	28.4	60.2	72.1
GC	0.034	0.031	0.033	0.042	0.006	-	8.82	16.3	22.3	31.3	49.1	42.8	72.2	84
GCR	0.034	0.030	0.033	0.043	0.013	0.005	-	12.5	19.8	39.5	58.7	52.5	82	92.7
GK	0.035	0.032	0.032	0.041	0.006	0.002	0.006	-	8.43	40.7	61.2	53.2	78.6	90.7
GST	0.031	0.027	0.030	0.033	0.005	0.004	0.004	0.003	-	43.3	63	53.6	77.4	88.4
CO	0.022	0.024	0.023	0.032	0.005	0.010	0.012	0.004	0.006	-	19.7	12.6	43.6	55.7
AR	0.025	0.022	0.023	0.027	0.002	0.008	0.014	0.005	0.005	0.001	-	8.33	32.7	42.8
BA	0.026	0.025	0.024	0.030	0.008	0.011	0.013	0.008	0.009	0.003	0	-	32.8	43.3
CL	0.033	0.031	0.029	0.034	0.015	0.021	0.030	0.020	0.026	0.011	0.012	0.016	-	13.1
FL	0.031	0.028	0.027	0.032	0.019	0.027	0.032	0.027	0.032	0.019	0.016	0.019	0.003	-

In the five independent simulations of the Bayesian clustering method analysis implemented in STRUCTURE (Pritchard *et al.* 2000), K = 4 was the most probable number of clusters found across the study area with average P (K|X) = -51614.08. However, the average value for K = 3 was very similar, P (K|X) = -51672.2, and was the point at which most explanatory power was gained (i.e. the K that captured most of the structure in the data). For K > 4 probability values decreased dramatically. A clear genetic differentiation between sampling sites located west and east of the Great Glen was already depicted for K = 2, therefore indicating that this was the major differentiation of the area. The differentiation of CL and FL from the other sampling sites became apparent at K = 3. Two very clear clusters were obtained, one comprising the four populations sampled west of the Great Glen and another comprising the most easterly sampled populations (CL-FL) (Figure 3.3). The other populations sampled comprised individuals with much higher levels of admixture but could be subdivided into two further groups, one including the estates AR-BA-CO-MA and another consisting of GC-GK-GCR-GST (Figure 3.3).

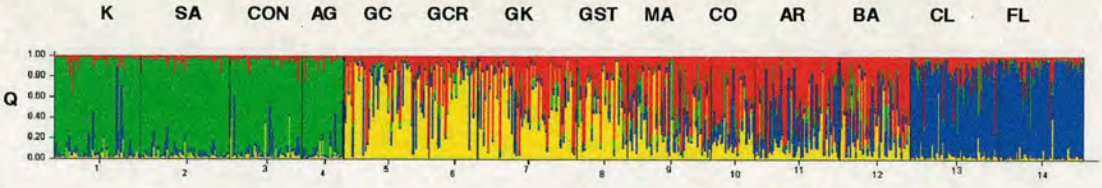


Fig. 3.3 Bayesian clustering analysis. Plot of the estimates of Q (estimated membership coefficient for each individual) for each cluster (K). The most probable number of genetic populations present in the data is K = 4. The vertical lines are broken into coloured segments showing the proportion of each individual assigned to each of the inferred K. Numbers and letters at the bottom and top of the graph correspond to the sampling estates. Test for K = 1-14 were performed using a 30,000 burnin period and 10^6 MCMC simulations in five independent runs.

3.4.3 Effect of geographical distance and landscape features on red deer population structure

3.4.3.1 Isolation by distance

The correlation between geographical distance and genetic differentiation for the whole study area was significant (r = 0.6, P = 0.0001), with 36.51 % of the genetic

differentiation explained by the model (Fig. 3.4). When the sample was divided into two geographical groups, west and east of the Great Glen, geographical distance explained 56.78 % of genetic differentiation in eastern subpopulations (r = 0.75, P = 0.0001) (Fig. 3.4). However, when not including the most distant estates sampled east of the Great Glen (FL-CL) only 32.14% of the genetic variance was explained by geographical distance (r = 0.57, P = 0.0169). In western populations, the correlation was very high but not significant (r = 0.97, P = 0.0811) (Fig.3.5) probably due to the smaller number of sampling sites west of the Great Glen as when conducting an analysis using four similarly distributed estates sampled east of Great Glen (MA, CO, GK, GST) a non-significant positive correlation was also found (r = 0.6977, P = 0.1217). When using only comparisons between populations at either side of the Great Glen, no correlation between genetic and geographical distance was found (r = -0.079, P = 0.63; Fig. 3.5).

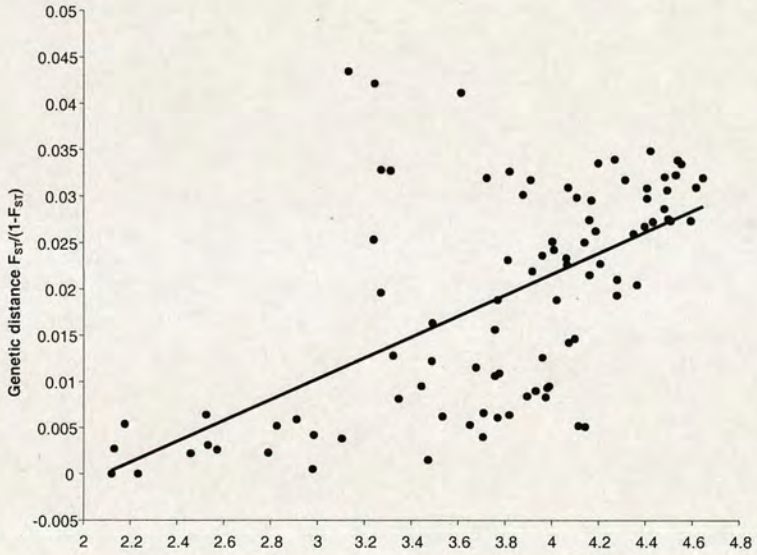


Fig. 3.4 Isolation by distance analysis. There was a positive correlation between geographical distance (ln km) and genetic distance ($F_{ST}/(1-F_{ST})$) between all pairs of sampling sites. Euclidean geographical distance explained 36.51% of the genetic variation in the study area (r = 0.6042, P < 0.01).

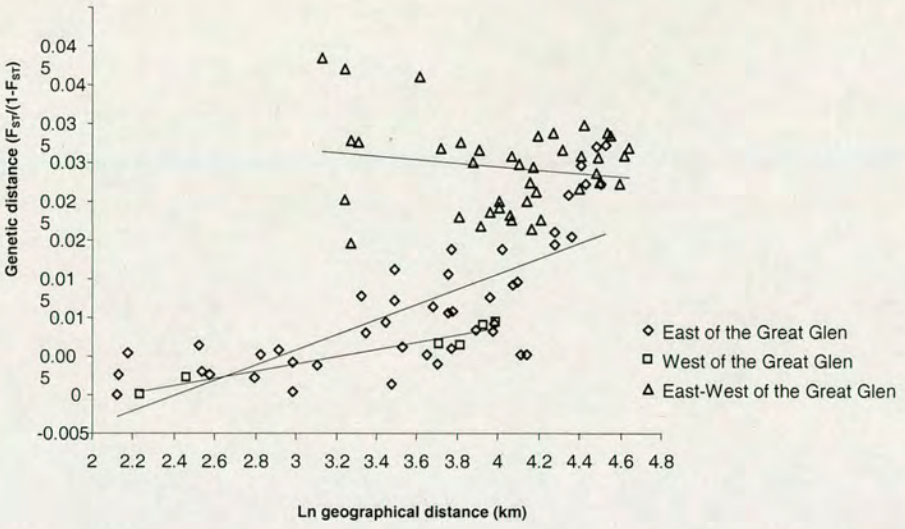


Fig. 3.5 Isolation by distance among populations sampled east of the Great Glen (r = 0.75, P = 0.0001), among populations sampled west of the Great Glen (r = 0.97, NS) and among populations sampled to either side of the Great Glen (r = 0.079, NS).

3.4.3.2 Location and strength of gene flow barriers

Figure 3.6 illustrates the location of the two most important barriers when using the program BARRIER. The first and major barrier was located along the Great Glen valley, with the barrier being strongest at Loch Linnhe and its strength diminishing northwards. The second barrier was located between the estates CL and BA.

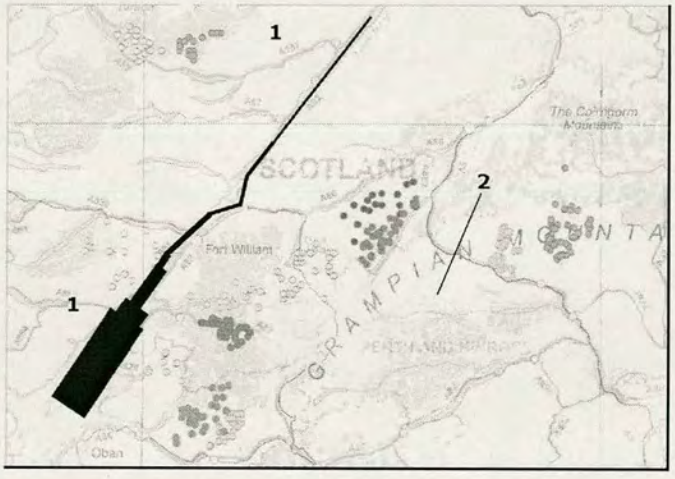


Fig. 3.6 Plot obtained from BARRIER showing the two first genetic barriers computed using mid point geographical locations for the populations and the residuals from a Mantel test between geographical and pair-wise F_{ST} genetic distances (to remove the effect of isolation by distance). Thickness of the lines illustrates the relative strength of the barrier.

3.4.3.3 Effect of landscape features on red deer population structure

The plots in figure 3.7 illustrate the results from simple Mantel tests between genetic distance and each of the landscape least-cost distance matrices. Inland lochs, railways and rivers were identified as potential red deer gene flow corridors as a higher amount of the genetic differentiation was explained when assigning cost values < 1. Sea lochs, roads, slopes and forests were identified as potential red deer gene flow barriers as cost values > 1 better explained the genetic differentiation. The cost value for which the highest r^2 was attained or at which r^2 reached a plateau was chosen as the cost value for each landscape feature to be used for subsequent analyses.

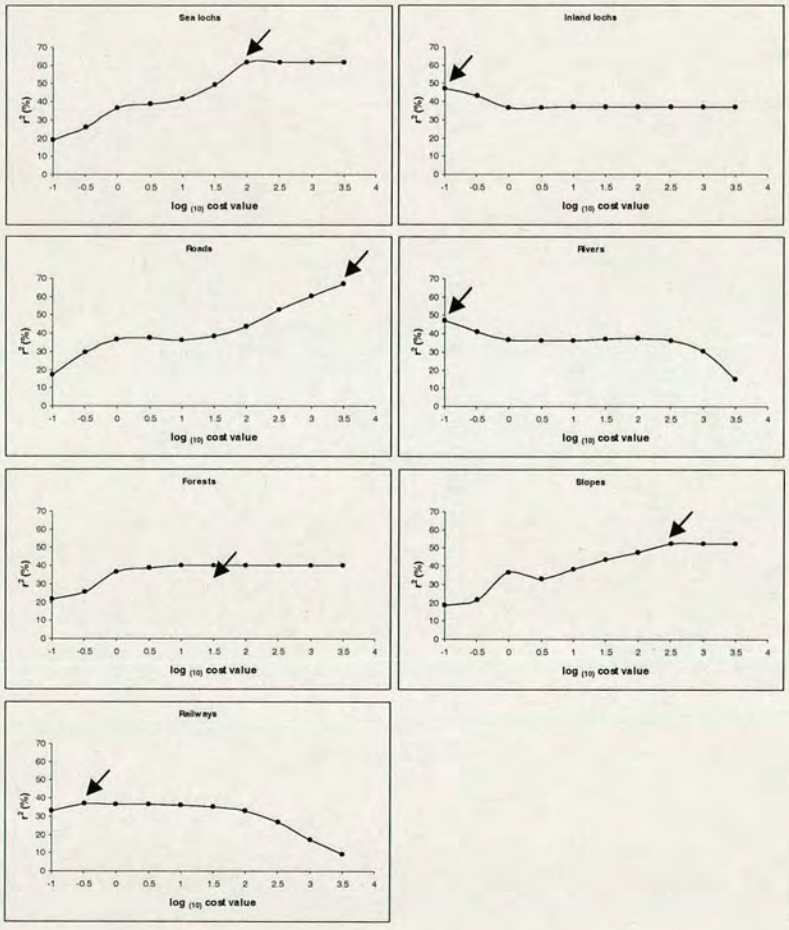


Fig. 3.7 Plots from simple Mantel tests analyses between genetic differentiation matrices against each of the landscape feature cost values. Cost values have been log-transformed for illustration purposes. r^2 (%) represents the amount of genetic differentiation explained by each landscape feature least-cost distance matrix using a particular cost value. All simple Mantel tests were significant (P = 0.0001). Black arrows indicate the cost values which explained the highest genetic differentiation or the first cost value at which r^2 plateaus.

Partial Mantel tests between pair-wise F_{ST} and each landscape feature correcting for the effect of geographical distance indicated that all landscape features assessed, except railways, had a significant effect on red deer gene flow (Table 3.4). Sea lochs and roads explained the greatest amount of genetic differentiation, 42.74 % and 30.77% respectively. The interaction of different landscape features on red deer gene flow was clearly indicated in the partial Mantel tests that included two landscape variables as the amount of genetic variation explained by a particular landscape feature varied depending on the order to which landscape variables were fitted in the model (Table 3.5).

Table 3.4 Results from partial Mantel tests between genetic differentiation and landscape distance after correcting for Euclidean distance between populations. All values had P < 0.001 except L (%) for railways that was not significant.

Landscape feature	D (%)	L (%)	R^2 (%)
Sea lochs	36.75	42.74	79.77
Roads	36.75	30.77	67.18
Mountain slopes	36.75	17.95	54.15
Forests	36.75	13.68	50.37
Rivers	36.75	12.82	48.8
Inland lochs	36.75	11.97	44.13
Railways	36.75	1	37.5

D (%): Percentage of the genetic variance explained by Euclidean distance

L (%): Percentage of the genetic variance explained by a landscape feature after removing the effect of distance

R^2 (%): Percentage of the variance explained by the model (distance + landscape feature)

Table 3.5 Results from partial Mantel tests including geographical distance and two landscape variables as explanatory variables to illustrate the interaction of different landscape features on the genetic differentiation between populations.

Landscape Feature (L1)	Landscape Feature (L2)	L1-D (%)	P-value	L2-L1-D (%)	P-value	(%) R^2
Sea Lochs	Roads	42.74	0.0001	3.41	0.0773	83.15
Roads	Sea Lochs	30.77	0.0001	16.24	0.0002	83.15
Sea Lochs	Forest	42.74	0.0001	2.56	0.1432	82.22
Forest	Sea Lochs	13.68	0.0001	31.62	0.0003	82.22
Sea Lochs	Slope	42.74	0.0001	1.71	0.2043	81.54
Slope	Sea Lochs	17.95	0.0001	27.35	0.0001	81.54
Sea Lochs	Rivers	42.74	0.0001	0	0.595	80.1
Rivers	Sea Lochs	12.82	0.0001	30.77	0.0009	80.1
Sea Lochs	Inland Lochs	42.74	0.0001	0	0.0007	79.96
Inland Lochs	Sea Lochs	11.97	0.0007	30.77	0.0001	79.96
Roads	Forests	30.77	0.0001	2.56	0.1419	69.67
Forests	Roads	13.68	0.0001	19.66	0.0005	69.67
Roads	Inland Lochs	30.77	0.0001	2.56	0.07	69.47
Inland Lochs	Roads	12	0.0007	20.51	0.0001	69.47
Roads	Slope	30.77	0.0001	2.56	0.1664	69.36
Slope	Roads	17.95	0.0001	15.38	0.0001	69.36
Roads	Rivers	30.77	0.0001	0	0.7659	67.28
Rivers	Roads	12.82	0.0004	17.95	0.0001	67.28
Forests	Inland Lochs	13.68	0.0002	5.13	0.0288	59.78
Inland Lochs	Forests	12	0.0003	11.11	0.0019	59.78
Slopes	Inland Lochs	17.95	0.0002	4.27	0.0621	58.15
Inland Lochs	Slopes	12	0.0006	9.4	0.0031	58.15
Slopes	Forests	17.95	0.0002	1.71	0.2341	55.69
Forests	Slopes	13.68	0.0001	0	0.0288	55.69
Rivers	Inland Lochs	12.82	0.0004	5.98	0.0139	55.43
Inland Lochs	Rivers	12	0.0004	6.84	0.014	55.43
Slopes	Rivers	17.95	0.0001	0	0.8708	54.15
Rivers	Slope	12.82	0.0006	5.13	0.0314	54.15
Forests	Rivers	13.68	0.0003	3.42	0.0086	53.71
Rivers	Forests	12.82	0.0005	4.27	0.0425	53.71

L1: First landscape feature fitted into the model after correcting for the effect of geographical distance

L2: Second landscape feature fitted into the model after correcting for the effect of geographical distance and the effect of landscape feature 1

L1-D (%): Percentage of the genetic variance explained by landscape feature 1 after correcting for geographical distance

L2-L1-D (%): Percentage of the genetic variance explained by landscape feature 2 after correcting for geographical distance and landscape feature 1

R^2 (%): Percentage of the genetic variance explained by the model. All values had P = 0.0001

3.5 Discussion

Red deer populations in mainland Scotland have been heavily affected by man for the last 300 years (Lowe & Gardiner 1974; Clutton-Brock *et al.* 1989; Hartl *et al.* 2003), offering ample opportunity for population bottlenecks during periods of intense hunting and translocation events. However, red deer populations within the study area had high genetic diversity, were in Hardy-Weinberg equilibrium and showed no evidence for genetic bottlenecks which would suggest that population declines during the past centuries were probably not severe enough to affect the effective population size.

Despite the relatively small scale of the study area, significant population genetic structure was found for Scottish Highland red deer ($F_{ST} = 0.019$) with most of the sites, except for those adjacent or closely located, being genetically differentiated from each other (Table 3.3). The degree of population differentiation found in this study is comparable to that found for Scottish red deer in the island of Rum despite studies were conducted at different geographical scales (Nussey *et al.* 2005; Nussey *et al.* in press). Although the global value of F_{ST} found for Scottish Highland red deer was low, clear sub-structuring in the area was found when using the program STRUCTURE (Fig. 3.3) for which substantial differences in allele frequencies are needed in order to define clusters (Pritchard *et al.* 2000). In fact, the levels of population differentiation might be underestimated in terms of F_{ST} values because of the high polymorphism of some of the markers used in this study (Hedrick 1999; Balloux & Lugon-Moulin 2002). This was confirmed by calculating the standardised measure of population differentiation proposed by Hedrick (2005) $G_{ST}' = 0.084$ which suggested higher levels of population genetic structure in the study area than those obtained with F_{ST}.

The genetic sub-structuring in the study area was mainly due to differentiation between estates sampled east and west of the Great Glen, followed by a further differentiation of CL- FL from other estates located east of the Great Glen (Fig. 3.1 &

3.3). Population differentiation in the central part of the study area was less marked, with individuals presenting much more admixed ancestries (Fig. 3.3). The lower sub-structuring in this central area could be due to the fact that we had a more continuous sampling effort in this area and so we obtained a gradation of allele frequencies making the delimitation of clusters by STRUCTURE more difficult (Pritchard *et al.* 2000). However, subsequent analyses on the effect of landscape features on red deer gene flow indicated that the lower sub-structuring might indeed reflect a higher movement of individuals between these central estates (discussed below). At a local scale, most of the adjacent or very closely located estates were not significantly genetically differentiated from each other (Table 3.3), reflecting a high gene flow between nearby areas and therefore agreeing with the common short dispersal distances of a few kilometres recorded in mark-recapture and GPS-tracking of red deer in mainland Scotland (Daniels & McClean 2003; Sibbald *et al.* unpublished data).

A significant pattern of isolation by distance was found across the study area and between estates sampled east of the Great Glen (Fig. 3.4, 3.5). For estates sampled west of the Great Glen a high correlation between genetic and geographical distance was found but was not significant probably to the small number of estates sampled in this area. No correlation between genetic and geographical distance was found when the analysis was performed using only comparisons between estates at either side of the Great Glen, therefore indicating that the Great Glen is a red deer gene flow barrier (Fig. 3.5). Isolation by distance despite the presence of gene flow barriers has been previously demonstrated (Bossart & Prowell 1998; Keyghobadi *et al.* 1999; Lugon-Moulin & Hausser 2002), indicating how important it is to analyse the effects of distance on population differentiation at different scales within a particular study area. The location of a gene flow barrier in the Great Glen was confirmed when using the program BARRIER (Fig. 3.6). The major genetic discontinuity was located at Loch Linnhe with its strength diminishing northwards. Additionally, a less severe barrier was detected between the estates of CL and BA confirming the differentiation of CL- FL from the rest of the estates sampled east of the Great Glen.

The genetic sub-structuring across the study area was only partly explained by the geographical distance between populations. Landscape features were shown to play an important role in red deer population differentiation as a greater proportion of the genetic variation was explained when adding landscape features to the distance matrices. This finding supports the current red deer management practice of delimiting Deer Management Group areas by considering natural or man-made landscape features that might surround a biological population and avoiding further division of the groups within such areas.

Although largely ignored in the current landscape genetics literature, we have shown in this study that choosing arbitrary cost values to construct least-cost distance matrices can have significant effects when identifying and quantifying the effect of landscape features on gene flow. Trying a range of cost values allowed us to identify which landscape features were potentially corridors or barriers to red deer gene flow and to decide which cost value was more appropriate to apply for each landscape feature. Therefore, we recommend that the evaluation of the performance of a range of cost values to be used to construct least-cost distance matrices should routinely be taken in similar landscape genetics studies.

Results from the partial Mantel tests, including two different landscape features as explanatory variables, indicated that when measuring the effect of a single landscape feature on the population genetic differentiation we might be capturing the effect of other landscape features present in the area. For example, sea lochs and roads explained the highest percentage of the genetic variance however, the amount of variance explained by each of the landscape features varied depending on the order to which they were fitted in the partial Mantel test (Table 3.5). Therefore, the intricate interaction of landscape features in nature should be taken into account to understand the effect of landscape on red deer gene flow.

Rivers and inland lochs were identified as biological corridors for red deer gene flow in the Scottish Highlands. In the study area, inland lochs are long but narrow and rivers are < 100 m wide (see Fig. 3.3) and therefore the cost of swimming lochs and rivers might be smaller than walking around them. Red deer swimming across inland lochs is a behaviour that it is frequently observed in the field (deer stalkers, personal communication). Although lakes and rivers have been found to be dispersal barriers for other mammals with high vagility, the dimensions of these landscape features and the scale of the study area were much larger than in our study (e.g. The Great Lakes in Wisconsin (USA) for wolf (*Canis lupus*), Mladenoff *et al.* 1995; The Peace river in western Canada for reindeer (*Rangifer tarandus*), McLoughlin *et al.* 2004). Lakes and rivers however, were not barriers for wolf dispersal in northern Poland (Jedrzejewski *et al.* 2004) and White *et al.* (2004) found that although bear dispersal decreased across the Mississippi river (1600 m wide) gene flow did not decrease across the White river (200 m wide).

Sea lochs, roads, mountain slopes, and forests were identified as red deer gene flow barriers in the study area. The strongest gene flow barriers were sea lochs; in particular Loch Linnhe appears to be the landscape feature responsible for the major genetic sub-structuring in the study area. This sea loch is 56 km long and the width varies between c. 500 m to c.15 km where it joins the Firth of Lorn. The low sea temperatures, currents and salty water might make the crossing of sea lochs highly costly for red deer and in our study area can be avoid by taking alternative land dispersal routes. Furthermore, other landscape features such as steep areas, fenced forests and roads adjacent to the shores of Loch Linnhe might reinforce the effects of this sea loch as a dispersal barrier (discussed below, Fig. 3.1 & 3.3). Sea water has also been identified as a barrier for dispersal for other mammals with high dispersal capabilities, for example the strait of Gibraltar has been suggested as gene flow barrier for wolf and jackal (*Canis aureus*) (Castella *et al.* 2000). Gene flow of Alaskan brown bear (*Ursus arctos*) was found to be reduced or sometimes absent between four insular populations when separated by stretches of sea water of a few kilometres; but continuous gene flow between populations was detected when the stretches of sea water were much narrower c. 600 m (Paetkau *et al.* 1998). Narrow

stretches of sea water in the study area such as Loch Etive and Loch Leven probably do influence the differentiation of the two pairs of closely related estates which they divide, GC-MA and GCR-GK (Fig. 3.1), as unlike the study of insular brown bear populations, alternative land dispersal routes exist in our study area which might be preferred for red deer movements.

Roads were also identified as gene flow barriers to red deer. In our study area the A9 road with an average of c. 7000 cars/day is the major tourist and trade route in Scotland (data from the Scottish Transport Statistics from 1998, http://www.scotland.gov.co.uk) and could potentially be responsible for the differentiation of the most easterly sampled estates (CL- FL). The other main road in the study area, the A82, another major tourist route in West Scotland (average c. 4000 cars/day) plus the adjacent A86, are also probably diminishing the exchange of red deer across the Great Glen. Negative effects of roads and/or railways on dispersal have also been reported for other mammals with high dispersal capabilities (e.g. Roach *et al.* 2001; Jedrzejewski *et al.* 2004; Epps *et al.* 2005; Waller & Servheen 2005; Riley *et al.* 2006). Railways though seem to not have any effect of the red deer population genetic differentiation. The Central Highland railway running parallel to the main road A9 could possibly be having an effect on red deer gene flow however; as the main genetic differentiation in the study area is along the Great Glen where no railway lines are found we might not be detecting its effect. Additionally, the West Highland railway seems to have little effect on gene flow as estates on either side of the railway showed high levels of gene flow (see Fig. 3.2 & 3.3).

The effect of roads on red deer gene flow however, needs to be explained in conjunction with the effect of other landscape features such as steep mountain slopes and forests adjacent to them. Roads are built along flatter areas corresponding to mountain valleys (see Fig. 3.3) and therefore steeper terrains on either side of roads may be a hindrance to the movement of red deer, which might prefer to move along rather than across roads. Alexander & Waters (2000) showed that in Banff National Park (Canada) optimal movement of red deer (elk), marten (*Martes americana*), wolf, bobcat and cougar (*Felis concolor*) was through low topographic complexity

areas and slopes lower than 5°. In our study we also observed a similar trend as populations sampled in the central part of the study area where terrain is less steep showed much less genetic differentiation (see Fig. 3.2 & 3.3).

Although forested areas have been found to influence connectivity between populations of other deer species (e.g. roe deer, Coulon *et al.* 2004; white-tailed deer, Long *et al.* 2005), in our study area forests were found to act as barriers for red deer gene flow. This might be explained by the fact that many forests in Scotland are fenced in order to avoid grazing and damaging of wooded areas by deer (Clutton-Brock & Albon 1989) and thus might restrict the movement of red deer between areas separated by forests. Fenced forested areas in Spain have also been shown to affect red deer gene flow (Martínez *et al.* 2002). In addition to the effect of fences on red deer dispersal, deer culling regimes in forests are much higher than in open hill areas in order to keep densities low and thus even if fences are trespassed the risk of mortality is very high (Ratcliffe 1984). In our study area, forests are particularly abundant along roads and in areas adjacent to the Great Glen, which might diminish the exchange of individuals between estates to either side of the Great Glen. The virtual absence of big patches of forested areas across the central part of the study area (see Fig. 3.3) coupled with the much flatter terrain found in that area and the absence of main roads might facilitate the exchange of individuals between those central estates.

3.6 Conclusions

Despite the relatively small scale of our study area, significant population structure was found for red deer, a large mammal with high dispersal capabilities. Landscape features were shown to play an important role as barriers and corridors for dispersal and thus important factors in shaping population structure of red deer in the Scottish Highlands. Inland lochs and rivers were identified as dispersal corridors and sea lochs, roads, forests and mountain slopes as red deer gene flow barriers. This is the first study to our knowledge that has assessed qualitatively and quantitatively the effect of a higher number of landscape features (natural and man-made) in animal

gene flow and has raised concern about the arbitrariness of cost values applied to construct least-cost distance matrices. We also stress the importance of considering the interaction between landscape features when interpreting the effect of an individual landscape feature on population genetic structure.

3.7 Acknowledgements

All stalkers and deer managers from the estates of Ardgour, Ardverikie, Ben Alder, Clunes, Conaglen, Corrour, Forest Lodge, Glencoe, Glencreran, Glenkinglass, Glenstrae, Kintail, Mamore and South Affric are greatly thanked for the collection of samples. Will Goodall-Copestake is thanked for his advice on the development of the multiplex protocol and microsatellite scoring. Angela Sibbald and Russell Hooper are thanked for information on unpublished data from their GPS-based studies. Betty Duff and David Elston from BIOSS are thanked for their help on statistical analyses. Jesus Mavárez provided valuable discussions and comments that greatly improved this manuscript. This project was funded by The Macaulay Development Trust. SEERAD - Scottish Executive supported F.J.P.-B.

3.8 References

Alexander SM, Waters NM (2000) The effects of highway transportation corridors on wildlife: a case study of Banff National Park. *Transportation Research* **Part C 8**, 307-320.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**, 155-165.

Bossart J, Prowell D (1998) Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *Trends in Ecology & Evolution* **13**, 202-206.

Castella V, Ruedi M, Excoffier L, *et al.* (2000) Is the Gibraltar Strait a barrier to gene flow for the bat *Myotis myotis* (Chiroptera: Vespertilionidae). *Molecular Ecology* **9**, 1761-1772.

Clobert J, Danchin E, Dhont AA, Nichols JD (2001) *Dispersal* Oxford University Press, Oxford.

Clutton-Brock TH, Albon SD (1989) *Red deer in the Highlands* BSP Professional Books, Oxford.

Cornuet JM, Luikart G (1996) Description and evaluation of two tests for detecting recent bottlenecks. *Genetics* **144**, 2001-2014.

Coulon A, Cosson J-F, Angibault JMA, *et al.* (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular Ecology* **13**, 2841-2850.

Coulon A, Guillot G, Cosson J-F, *et al.* (2006) Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Molecular Ecology* **15**, 1669-1679.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution* **15**, 290-229.

Daniels M, McClean C (2003) Red deer calf tagging programmes in Scotland - an analysis. *Deer (The Journal of the British Deer Society)* **12**, 420-423.

Di Rienzo A, Garza JC, Valdés AM, Slatkin M, Friemer NB (1994) Mutational processes of simple sequence repeat loci in human populations. *Proceedings of the National Academy of Sciencies of the United States of America* **91**, 3166-3170.

Dobson FS (1998) Social structure and gene dynamics in mammals. *Journal of Mammalogy* **79**, 667-670.

Epps CW, Palsbøll PJ, Wehausen JD, *et al.* (2005) Highways block gene flow and cause a rapid decline in genetic diversity of desert bighorn sheep. *Ecology Letters* **8**, 1029–1038.

Fisher RA, Ford EB (1947) The spread of a gene in natural conditions in a colony of moth *Panaxia dominula* L. *Heredity* **1**, 143–174.

Funk CW, Blouin MS, Corn PS, *et al.* (2005) Population structure of Columbia spotted frogs (*Rana luiteventris*) is strongly affected by the landscape. *Molecular Ecology* **14**, 483-496.

Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology* **13**, 2481-2490.

Goudet J (1995) FSTAT version 1.2: a computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**, 359.

Hartl G, Zachos F, Nadlinger K (2003) Genetic diversity in European red deer (*Cervus elaphus* L.): anthropogenic influences on natural populations. *Comptes Rendus Biologies* **326**, 37-42.

Hartl GB, Zachos FE, Nadlinger K, *et al.* (2005) Allozyme and mitochondrial DNA analysis of French red deer (*Cervus elaphus*) populations: genetic structure and its implications for management and conservation. *Mammalian Biology* **70**, 24-34.

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**, 313-318.

Hedrick PW (2005) A standardised genetic differentiation measure *Evolution* **59**, 1633-1638.

Jedrzejewski W, Niedzialkowska M, Nowak S, Jedrzejewska B (2004) Habitat variables associated with wolf (*Canis lupus*) distribution and abundance in northern Poland. *Diversity and Distributions* **10**, 225-233.

Keyghobadi N, Roland J, Strobeck C (1999) Influence of landscape on the population genetic structure of the alpine butterfly Parnassius smintheus (Papilionidae). *Molecular Ecology* **8**, 1481-1495.

Long ES, Diefenbach DR, Rosenberry CS, Wallingford BD, Grund MD (2005) Forest cover influences dispersal distance of white-tailed deer. *Journal of Mammalogy* **86**, 623-629.

Lowe A, Harris S, Ashton P (2004) *Ecological Genetics. Design, Analysis, and Application* Blackwell Science Ltd, Oxford, UK.

Lowe VPW, Gardiner AS (1974) A re-examination of the subspecies of Red deer (*Cervus elaphus*) with particular reference to the stocks of Britain. *Journal of Zoology (London)* **174**, 185-201.

Lugon-Moulin N, Hausser J (2002) Phylogeographical structure, postglacial recolonization and barriers to gene flow in the distinctive Valais chromosome race of the common shrew (*Sorex araneus*). *Molecular Ecology* **11**, 785-794.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* **89**, 238 - 247.

Manel S, Schwartz M, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution* **18**, 189-197.

Manni F, Guérard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by "Monmonier's algorithm". *Human Biology* **76**, 173-190.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**, 209-220.

Marshall T, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639-655.

Martínez JG, Carranza J, Fernández-García JL, Sánchez-Prieto CB (2002) Genetic variation of red deer populations under hunting exploitation in Southwestern Spain. *Journal of Wildlife Management* **66**, 1273-1282.

McLoughlin PD, Paetkau D, Duda M, Boutin S (2004) Genetic diversity and relatedness of boreal caribou populations in western Canada. *Biological Conservation* **118**, 593-598.

McRae BH, Beier P, Dewald LE, Huynh LY, Keim P (2005) Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American puma. *Molecular Ecology* **14**, 1965-1977.

Mladenoff DJ, Sickley TA, Haight RG, Wydeven AP (1995) A regional landscape analysis and prediction of favorable gray wolf habitat in the northern Great Lakes region. *Conservation Biology* **9**, 279-294.

Moritz C (2002) Strategies to protect biological diversity. *Systematic Biology* **21**, 238-254.

Neigel JE (1997) A comparison of alternative strategies for estimating gene flow from genetic markers. *Annual Review of Ecology & Systematics* **28**, 105-128.

Nussey DH, Coltman DW, Coulson T *et al* (2005). Rapidly declining fine-scale spatial structure in female red deer. *Molecular Ecology* **14**, 3395-3405.

Nussey DH, Pemberton J, Donald A, Kruuk LEB (2006) Genetic consequences of human management in an introduced island population of red deer (*Cervus elaphus*). *Heredity*, in press.

Paetkau D, Shields GF, Strobeck C (1998) Gene flow between insular, coastal and interior populations of brown bears in Alaska. *Molecular Ecology* **7**, 1283-1292.

Park SDE (2001) *Trypanotolerance in West African cattle and the population genetic effects of selection*, Ph.D. thesis, University of Dublin.

Piertney SB, MacColl ADC, Bacon PJ, Dallas JF (1998) Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Molecular Ecology* **7**, 1645-1654.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**, 502-503.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Ratcliffe PR (1984) Population-dynamics of red deer (*Cervus elaphus* L.) in Scottish commercial forests. *Proceedings of the Royal Society of Edinburgh, section B -Biological Sciences* **82**, 291-302.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution* **49**, 1280-1283.

Riley SPD, Pollinger JP, Sauvajot RM, *et al.* (2006). A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology* **15**, 1733-1741.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Roach JL, Stapp P, Van Horne B, Antolin MF (2001) Genetic structure of a metapopulation of black-tailed prairie dogs. *Journal of Mammalogy* **82**, 946-959.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.

Ryman N, Jorde PE (2001) Statistical power when testing for genetic differentiation. *Molecular Ecology* **10**, 2361-2374.

Sacks BN, Brown SH, Ernest HB (2005) Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history. *Molecular Ecology* **13**, 1265-1275.

Scottish Transport Statistics website. http://www.scotland.gov.co.uk

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**, 264-279.

Spear SF, Peterson CR, Matocq MD, Storfer A (2005) Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Molecular Ecology* **14**, 2553-2564.

Spencer CC, Neigel JE, Leberg PL (2000) Experimental evaluation of the usefulness of microsatellite DNA for detecting bottlenecks. *Molecular Ecology* **9,** 1517-1528.

Sugg DW, Chesser RK, Dobson FS, Hoogland JL (1996) Population genetics meets behavioural ecology. *Trends in Ecology & Evolution* **11**, 338-342.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535-538.

Vignieri SN (2005) Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). *Molecular Ecology* **14**, 1925-1937.
Waller JS, Servheen C (2005) Effects of transportation infrastructure on grizzly bears in Northwestern Montana. *Journal of Wildlife Management* **69**, 985-1000.
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.
White TH, Bowman JL, Leopold BD, *et al.* (2000) Influence of Mississippi alluvial valley rivers on black bear movements and dispersal: implications for Louisiana black bear recovery. *Biological Conservation* **95**, 323-331.
Whittow J (1992) *Geology and Scenery in Britain* Chapman & Hall, London, UK.
Wright S (1943) Isolation by distance. *Genetics* **28**, 139-156.
Wright S (1978) *Evolution and the genetics of populations* University of Chicago Press, IL, Chicago.

Chapter 4

Dispersal is not strongly male-biased in red deer (*Cervus elaphus*) in the Scottish Highlands

S. Pérez-Espona,[*§] F.J. Pérez-Barbería, [§] C. D. Jiggins, [*] I.J. Gordon, † J. M. Pemberton[*]

*Institute of Evolutionary Biology, School of Biological Sciences,The University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

§ The Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK

† Sustainable Ecosystems, CSIRO – Davies Laboratory, PMB PO Aitkenvale, Qld 4814, Australia

Corresponding author:

Sílvia Pérez-Espona,

Institute of Evolutionary Biology, School of Biological Sciences, The University of Edinburgh

The King's Buildings, West Mains Road, Edinburgh EH9 3J, UK

Fax: ++ 44 (0)131 650 6564

Tel: ++44 (0)131 651 3612

E-mail: s.perez-espona@sms.ed.ac.uk

S.Pérez-Espona,[*§] F.J. Pérez-Barbería, [§] C. Jiggins, [*] I.J. Gordon, † J. Pemberton[*]

Author's contribution:

SPE organised the sampling, carried out the laboratory work, performed the statistical analysis and wrote the manuscript. The project was set up by FJPB and IJG. The whole project was supervised by FJPB, JMP, CDJ and IJG. All authors critically read and improved drafts of this manuscript.

4.1 Abstract

For mammals with a polygamous mating system, dispersal is predicted to be male-biased. However, discrepancies in the direction and the extent of the bias in dispersal for a particular species are rising with the increase in empirical studies. Studies on sex-specific dispersal patterns in large mammals with high dispersal capabilities are rare due to the difficulties associated with measuring dispersal in the field. Indirect methods using molecular markers to estimate sex-biased dispersal are proving successful in providing insights on sex-specific dispersal patterns for a wide range of organisms for which field-based studies are impractical. In this study the current direction and extent of sex-biased dispersal of a large polygynous mammal, the red deer (*Cervus elaphus*), was assessed in thirteen populations from the Scottish Highlands by comparing 21 bi-parentally inherited microsatellite markers between male and female post-dispersal individuals. Estimates of F_{ST} and relatedness indicated weak male-biased dispersal in the study area. However, individual-based spatial autocorrelation analysis revealed that male-biased dispersal is weak even at small geographical distances (0-5 km) at which differences in dispersal between sexes have been shown to be skewed in field-based studies. Demographic, environmental and anthropogenic factors that might be influencing red deer sex-specific dispersal patterns in the Scottish Highlands are discussed. Male red deer behaviour such as males returning to natal or local areas to breed is also a factor that could explain the limited extent of male-biased dispersal detected in this study.

Keywords:

Cervus elaphus, microsatellites, sex-biased dispersal, red deer

4.2 Introduction

Sex-biased dispersal, the asymmetric dispersal of sexes through which one sex remains in or returns to the area where it was born to breed (philopatry), has been a subject of discussion during the past few decades as it is a profound difference in the life history of males and females in some vertebrates (e.g. Pusey & Packer 1987; Clutton-Brock & Albon 1989; Johnson & Gaines 1990; Waser *et al.* 1994). Sex-specific dispersal patterns can strongly influence the spatial genetic structure of populations (Chepko-Sade & Halpin 1987; Stenseth & Lidicker 1992) and, therefore, affect the evolutionary trajectory of populations and species. To study the direction and the extent to which sex-biased dispersal occurs in nature can not only provide insights into the evolutionary biology of a particular species, but also provide information on specific ecological requirements of both sexes, a pre-requisite for many conservation and management programs (e.g. Maehr *et al.* 2002; Zenger *et al.* 2003; Rubin & Bleich 2005).

The direction of sex-biased dispersal observed in nature has been tightly linked to species-specific breeding systems. As a general rule, for mammals, in which the majority of species have polygamous or promiscuous mating systems, male-biased dispersal is observed, with males dispersing at higher rates and/or larger distances than females and most females remaining in the area where they were born (Greenwood 1980; Dobson 1982). In contrast, in birds, for which monogamy is more common, female biased dispersal is expected. Although the evolutionary causes of sex-biased dispersal are thought to be multiple and are not yet fully understood (Gandon & Michalakis 2001), three main but not mutually exclusive hypotheses are widely accepted: local resource competition, local mate competition and inbreeding avoidance. For polygynous mammals, under the local resource competition hypothesis, females benefit from being philopatric because they depend on exploiting local resources in order to provide all parental care (Johnson & Gaines 1990). In contrast, males do not benefit from being philopatric because they do not defend local resources in order to attract females and parental care is generally minimal or absent (Greenwood 1980; Clutton-Brock & Albon 1989). However, there are

exceptions as males of some mammal species defend territories, e.g. roe deer *Capreolus capreolus* (Wahlstrom & Liberg 1998). The local mate competition hypothesis predicts that for polygynous mammals in which males mate with multiple females, competition between males for mating opportunities with females would drive males to disperse so as not to compete with relatives (Greenwood 1980; Dobson 1982). The inbreeding avoidance hypothesis suggests a high risk of inbreeding depression if mating occurs between relatives, which can be avoided if one of the sexes disperses (Johnson & Gaines 1990; Wolff 1993; Perrin & Goudet 2001). In polygynous mammals, males are less affected by inbreeding depression as they mate with multiple females but for females, which normally mate once a year, the risk of mating with a relative is very costly; so females would benefit if they disperse (Greenwood 1980). However, if females choose to mate with unrelated males (mate choice), males then would need to disperse in order to appear unfamiliar to females (Perrin & Goudet 2001). Based on game theory and using a kin-selection mathematical model, Perrin & Mazalov (2000) showed that for mammals with polygynous mating systems, male-biased dispersal would be expected if competition for mates between males exceeds the local resource competition among females.

However, as the number of empirical studies on sex-specific patterns of dispersal increases, discrepancies in the predicted direction and the extent of sex-biased dispersal are rising. For example, for Eurasian badgers (*Meles meles)* several studies have reported male-biased dispersal (e.g. Kruuk 1989; Rogers *et al.* 1998; Roper *et al.* 2003), others female-biased dispersal (Christian 1994; Woodroffe *et al.* 1995; Tuyttens *et al.* 1995) and a recent molecular study including eight British populations of badgers did not find overall significant sex-biased dispersal (Pope *et al.* 2006). Similarly, studies on the European rabbit *Oryctolagus cuniculus* in Australia have shown males dispersing more than females in only some populations, with levels of philopatry in young rabbits ranging from 18-90% for males and 32-95% for females depending on the population (Richardson *et al.* 2002).

Some of the discrepancies in sex-biased dispersal estimates between studies could be due to limitations of the methodology used to measure dispersal and/or the scale at

which dispersal has been measured in different studies. The limitations of field-based methods such as the difficulties of marking individuals, the small number of individuals that can be followed and the underestimation of long distance dispersal events have already been extensively reported (e.g. Slatkin 1985; Koening 1996). Indirect methods using molecular markers, although not completely free of disadvantages, have been shown to overcome some of these difficulties associated with field-based studies and therefore provide the potential to study sex-biased dispersal for species for which field-based studies are absent or impractical (Goudet *et al.* 2002; reviewed in Prugnolle & de Meeus 2002). Unlike the measures of mobility obtained by direct methods, indirect methods provide a measure of effective dispersal (dispersal and subsequent reproduction) (reviewed in Prugnolle & de Meeus 2002). Indirect methods, though, provide a temporal average measure of dispersal rather than current dispersal and can be influenced by several historical processes (Slatkin 1987) and thus it could be difficult to make comparisons between estimates obtained by indirect and direct methods. However, indirect methods involving comparisons of bi-parentally inherited markers such as microsatellites between sexes can provide a measurement of current dispersal similar to those obtained by field-based studies if the analysis only comprises post-dispersal individuals and if their offspring are not included in the sample (since the offspring will inherit a random set of alleles from their mother and father and thus the sex-biased dispersal signature will be lost in one generation) (reviewed in Prugnolle & de Meeus 2002; Goudet *et al.* 2002).

Discrepancies in the direction or extent of sex-biased dispersal between studies can not only be due to differences in the methodologies used among studies, but can also arise due to the spatial and temporal heterogeneity of demographic and environmental factors influencing the costs and benefits of sex-specific dispersal in different populations (Boujemadi *et al.* 1999; Whitlock 2001; Wiens 2001). Furthermore, the influence of human management on economically important species can also influence differences in dispersal patterns when compared to that of undisturbed populations (Tuyttens & MacDonald 2000). Therefore, the study of sex-biased dispersal in different populations of a particular species living in different

environmental conditions will help elucidate the evolutionary and ecological factors that might make one sex more prone to dispersing than the other.

The red deer (*Cervus elaphus*) is a mammal with a strongly polygynous mating system, characterised by female philopatry and male-biased dispersal (Clutton-Brock *et al.* 1982a; Clutton-Brock *et al.*1989). Red deer females rarely disperse and often stay in the area where they were born forming groups of related females, matrilines (Clutton-Brock *et al.* 1982a, Clutton-Brock & Albon 1989; Albon *et al.* 1992). In contrast, males stay within the female groups until the juvenile stage and then, when they are between 2 and 4 years old, leave the natal area in search of a new home range (Clutton-Brock *et al.* 1982a; Clutton-Brock & Albon 1989).

In mainland Scotland, red deer are abundant, with a population of c. 400,000 occupying three quarters of the Scottish Highlands. The population has gradually doubled during the past 40 years, probably due to the reduction in sheep stocks and changes in weather, with milder winters and hotter summers (Clutton-Brock & Albon 1989; Clutton-Brock & Albon 1992; Clutton-Brock *et al.* 2004). Deer stalking is still an important land use in the Scottish Highlands and populations of red deer have traditionally been maintained at high densities in order to provide enough males due to their high value as hunting trophies (Clutton-Brock & Albon 1989; Clutton-Brock *et al.* 2004). In particular, densities of females in Highland Scotland are maintained high by estates in order to attract males from neighbouring estates or keep males resident on particular estates (Clutton-Brock & Albon 1989; Clutton-Brock *et al.* 2002).

Despite the fact that a large amount of research on different aspects of red deer dispersal has been conducted in the long-term study area of the North Block in the island of Rum (Scotland, UK) (e.g. Clutton-Brock *et al.* 1982a; Pemberton *et al.* 1992; Coulson *et al.* 1997; Clutton-Brock *et al.* 2002; Nussey *et al.* 2005; Nussey *et al.* in press), information on red deer dispersal in mainland Scotland, despite representing a "more natural population", is limited to sparse recovery data from calf tagging programs (Red Deer Commission 1978, 1982, 1983, 1989; Daniels &

McClean 2003), a study examining dispersal during the mating season by following 18 adult red deer carrying Global Positioning Satellite (GPS) collars (Sibbald *et al.* unpublished data), and a recent study assessing the effects of landscape features on Scottish Highland red deer using microsatellite markers (See Chapter 3 or Pérez-Espona *et al.* submitted).

In the North Block study area on Rum where the red deer herd has been released from culling since 1973, a study examining fine-scale genetic structure using microsatellite markers and census data across a 24-year period showed that genetic relatedness among females was significantly higher than males for distances < 500 m, and that with average dispersal rates out of the study areas were 11.7% and 30% for adult females and males, respectively (Nussey *et al.* 2005). However, when examining dispersal patterns among the different subdivisions of the study area, discrepancies in the bias of the dispersal were found, with some areas showing male-biased dispersal and others female-biased dispersal (Nussey *et al.* 2005). Analysis of red deer calf tagging programmes in the Scottish mainland showed that on average males dispersed larger distances than females, with dispersal distances ranging from 3.3-7.4 km and from 1.9 –3.5 km for males and females, respectively (Daniels & McClean 2003). However, variation in the distances dispersed by males and females could be observed within the different studied areas, with some areas showing male-biased dispersal and others not showing any sex-biased dispersal (Daniels & McClean 2003). Analyses from these mainland calf tagging programs suggest that sex-specific dispersal patterns of red deer in the Scottish mainland are likely to differ from those found on the Island of Rum and therefore further investigation is required. Information on sex-biased dispersal of red deer on the mainland of Scotland will not only provide insights into the evolutionary and ecological aspects of red deer dispersal but also provide helpful information for current management programs.

In this study we used an indirect molecular-based approach to assess the direction and the extent of current sex-biased dispersal in red deer (*Cervus elaphus*) in the Scottish mainland. In addition, we used an individual-based spatial autocorrelation

analysis to assess at which distances the extent of a potential sex-biased dispersal was more prominent.

4.3 Materials and Methods

4.3.1 Sample collection and DNA extraction

Samples for this study consisted of either an ear tip or jaw of 568 adult wild red deer, 318 males and 250 females, aged 4 years or older as long-term studies of red deer on the North Block in the island of Rum have shown that juveniles usually leave the natal area when they are 2-4 years old (Clutton-Brock & Albon 1989). Samples were collected from legally shot individuals during the 2003-2004 hunting (culling) season in 13 estates in the Scottish Highlands (Fig. 4.1). Male samples were collected during the male culling season (1st July-20th October – Deer Commission for Scotland website) with most of the samples obtained at the end of summer. Female samples were collected during the female culling season (21st October-15th February – Deer Commission for Scotland website) with most of the samples collected in winter. With this sampling strategy we maximised the potential of detecting any bias in dispersal as at the end of summer males are supposed to reach their maximal dispersal rates and/or distances in order to find females with which to mate during the rutting season (October – November) (Clutton-Brock *et al.* 1982a; Clutton-Brock & Albon 1989; Sibbald *et al.* unpublished data).

Tissue samples were stored either in a -20ºC freezer or in tubes containing 100% ethanol. Genomic DNA was extracted from ear or jaw muscle using the DNAace Spin Tissue Mini Kit (Bioline) or with the DNEasy Tissue Kit^{TM} (QIAGEN), following the manufacturer's instructions.

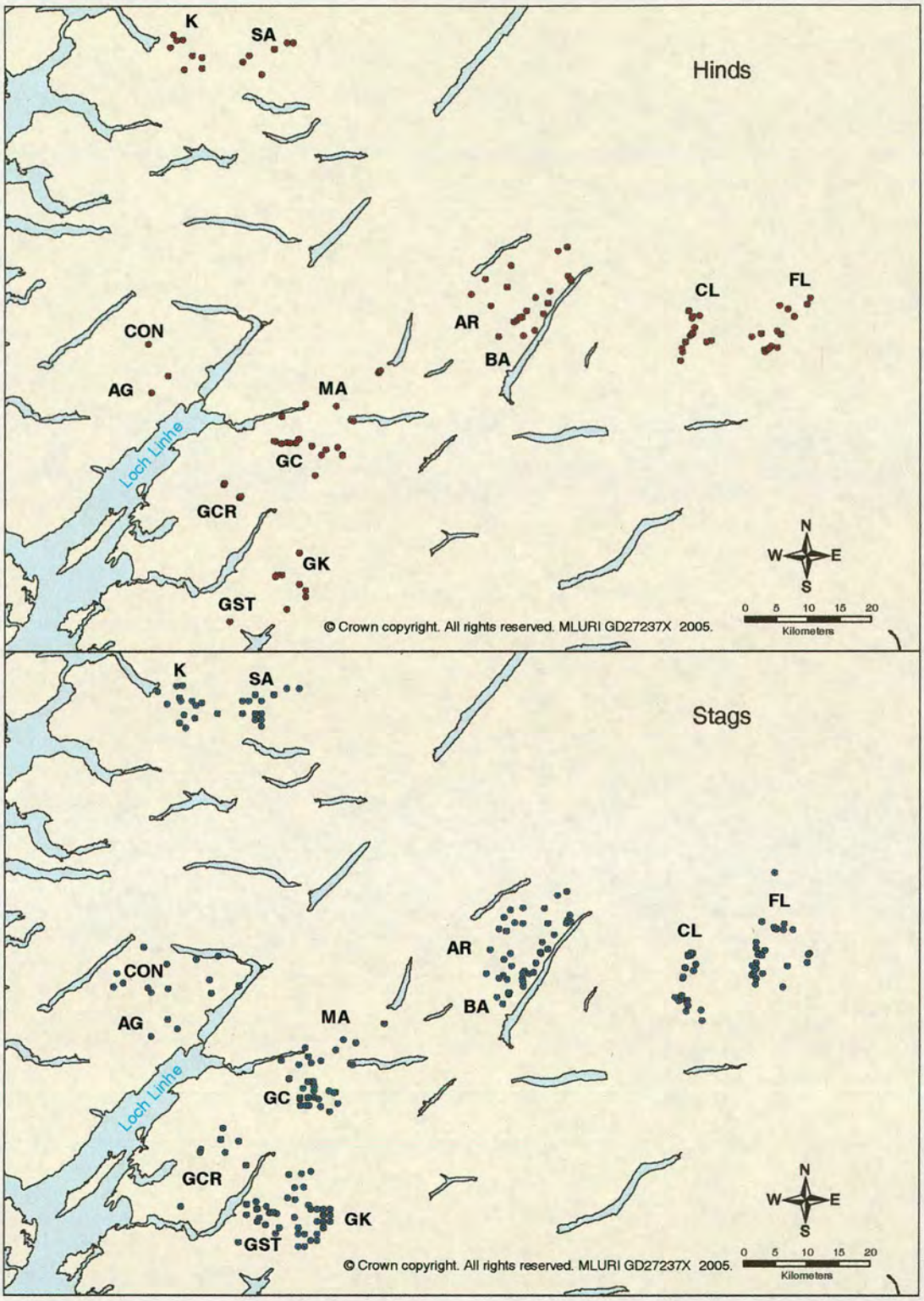


Figure 4.1. Map of the study area showing sampling locations for females (above) and males (below). Dots represent actual locations rather than number of individuals (i.e. several individuals were sampled in the same location). Abbreviations correspond to the name of estates from which samples were provided: **FL** = Forest Lodge (n_m=29, n_f =23), **CL**= Clunes(n_m=24, n_f =26), **BA**= Ben Alder(n_m=30, n_f =18), **AR**= Ardverikie(n_m=30, n_f =22), **MA**= Mamore(n_m=24, n_f =19), **GC**= Glencoe, (n_m=23, n_f =23) **GCR**= Glencreran(n_m=19, n_f =11), **GK**= Glenkinglass(n_m=35, n_f =14), **GST**= Glenstrae(n_m=23, n_f =5), **CON**= Conaglen(n_m=30, n_f =22), **AG**= Ardgour(n_m=30, n_f =18), **SA**= South Glen Affric(n_m=24, n_f =26), **K**= Kintail(n_m=29, n_f =18).

4.3.2 Sex determination

To verify that no mistakes regarding the gender of the sample occurred when labelling the sample in the field or in the lab we run all samples for a sex marker. The sex marker consisted of the amplification of an intron within the zinc finger gene of the X and Y chromosomes (Shaw *et al.* 2003). Amplification of the Zfx and Zfy introns was done using the primers LGL331 and LGL335 (Cathey *et al.* 1998). PCR reactions and cycles were as described in Shaw *et al.* (2003). After amplification, samples were run on a 1.5% agarose gel at 100V for about 3 h. For females one band was observed in the gel and two bands were observed for males as the intron of the Y chromosome is smaller than the one in the X chromosome.

4.3.3 Microsatellite genotyping

A 21 microsatellite loci high-throughput protocol was developed for PCR and sequencer load multiplexing. Primer sequences, details of the 21-locus multiplex kit, polymerase chain reactions (PCRs), measures to check for consistent genotyping and genotyping success are described in Chapter 2 or in Pérez-Espona *et al.* (submitted). Multiplex PCR products were run on an ABI 3730 capillary sequencer (Applied Biosystems) together with the internal size standard GeneScan 500 LIZ (Applied Biosystems). Fragment analysis was done using the software GeneMapper™ version 3.0 (Applied Biosystems). Details about genetic diversity measures obtained with the multiplex kit are described in detail in Chapter 2 & 3.

4.4 Data analysis

4.4.1 Detection of sex-biased dispersal

Sex-biased dispersal detection tests comparing bi-parentally inherited microsatellite markers between males and females were calculated in FSTAT version 2.9.3. (Goudet 2002, modified from Goudet 1995) using five different statistics: F_{IS}, F_{ST},

relatedness, mAIc and vAIc. If red deer dispersal is male-biased, a positive and higher value of F_{IS} is expected for males than females as this statistic indicates how well the genotype frequencies match Hardy-Weinberg expectations (Hartl & Clark 1997). Therefore, if males are the predominantly dispersing sex, the males in a particular population will be a mixture of immigrant and resident individuals and due to the Wahlund effect a heterozygote deficiency and thus a positive F_{IS} will be observed. In contrast, we would expect higher F_{ST} values for females than males. F_{ST} measures the degree of differentiation between populations (Hartl & Clark 1997) so for the philopatric sex we would expect allele frequencies between populations to be more different than among the dispersing sex. Within-population relatedness (*r*) is also expected to be higher for females since if females tend to stay in the area where they were born they would be expected on average to be more related to each other than males. This statistic is connected to F_{ST} by the relationship $r = 2 F_{ST} /(1-F_{IT})$ (Queller & Goodnight 1989), with F_{IT} being a measure of inbreeding in individuals relative to the total population, which is expected to be identical for males and females (Goudet *et al.* 2002). The Assignment Index (AI) calculates the probability that an individual originated in the population where it has been sampled (Paetkau 1995; Favre *et al.* 1997). For multilocus data a corrected AI statistic is used, AIc (Goudet *et al.* 2002). The mean AIc (mAIc) is expected to be higher for females as for the philopatric sex, on average, more individuals will be assigned to the population where they were sampled. Variance of AIc (vAIc) is expected to be higher for males as the dispersing sex sampled at any site will contain a mixture of residents and immigrants.

Visualisation of differences in population genetic structuring between sexes was achieved by conducting a factorial correspondence analysis using the option "AFC sur populations" in GENETIX version 4.05.2 (Belkhir *et al.* 2004) using allelic data for each sex separately.

4.4.2 Sex bias in relation to spatial distance

Individuals used in this study for which sampling spatial coordinates were available (297 males and 240 females) were used to assess the existence and the extent of sex-biased dispersal at different geographical distance intervals using an individual-based spatial autocorrelation analysis approach. ARCGIS version 9 (ESRI) was used to plot individual coordinates on a map of the sampling area (Fig. 4.1). The distribution of inter-individual Euclidean spatial distances were plotted in order to assess the degree of sampling bias between sexes. Additionally, distances to the nearest neighbour were also calculated for each sex independently by using the free extension "Hawth's Analysis Tools" for ARCGIS (Beyer 2004). To detect if the bias due to sampling more females in a particular location in comparison to males was significant we compared pair-wise distances and nearest neighbour distances of males and females using an independent samples t-test in the software SPSS version 8 (SPSS Inc.).

The software SPAGEDI version 1.2 (Hardy & Vekemans 2002) was used to test for any association between genetic kinship and geographical distance. An inter-individual relative kinship coefficient (Ritland 1996) defined as Fij = (Qij –Qm)/ (1-Qm), with Qij being the probability of identity in state for random genes from individuals i and j, and Qm the probability of identity by state for genes coming from random individuals from the whole population sample was calculated (Vekemans & Hardy 2004). The association between inter-individual kinship values and spatial distance is obtained by taking an average of the pair-wise statistics for a number of distance intervals and subsequent regression of these values on spatial distances or their logarithm. The association between genetic distance and spatial distance was tested for the global pool of data and at 10 arbitrary distance intervals: 0-5, 5-10, 10-20, 20-30, 40-50, 50-60, 60-70, 70-80, 80-112 km. These distance intervals were chosen because all of them contained a high number of pair-wise comparisons (in our study > 1000 for females and >2000 for males), more than 50% of individuals were represented at least once in the interval, and the coefficient of variation of the number of times each individual was represented was ≤ 1, as recommended by SPAGEDI's authors in the software's manual. To assess the effect of sex-biased

dispersal at finer scale we performed an additional spatial autocorrelation analysis including only individuals separated by distances 0-5 km at 6 intervals (0-500, 500-1000, 1000-2000, 2000-3000, 3000-4000, 4000-5000m). However, due to the smaller number of individuals included in these short distance intervals, the conditions of > 50% of individuals represented at least once in the interval and coefficient of variation ≤ 1 were not always accomplished. However, the number of pairs of individuals was always >100 for males and females in all intervals, except for the interval 500-1000m for females where there were 59 pairs. Statistical significance for all average genetic distances at each of the geographical distance intervals (small and larger scale) were given as 95% confidence intervals for the null hypothesis (no association between genetic and geographical distance) after comparing the observed values to the frequency distributions of 10,000 random permutation tests of location, individual and loci.

4.5 Results

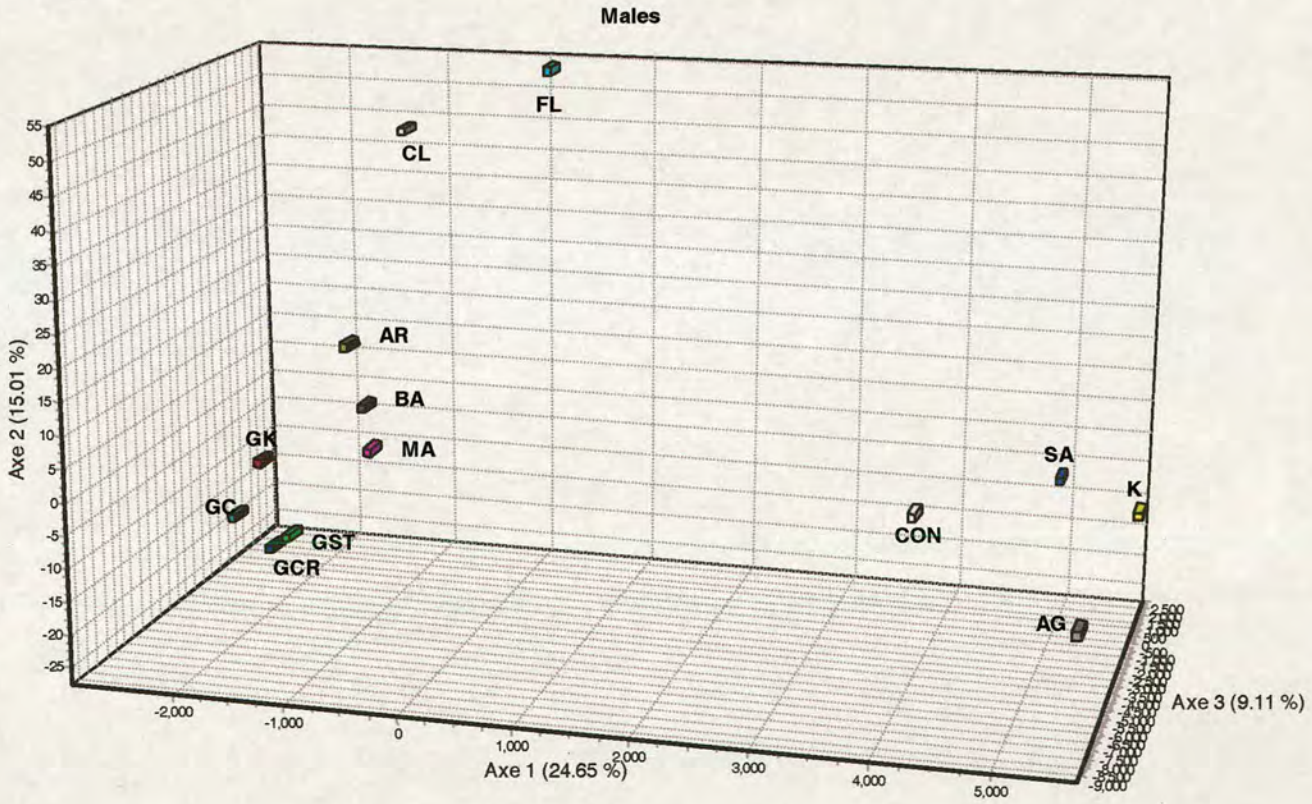
4.5.1 Detection of sex-biased dispersal

From the five tests undertaken in FSTAT 2.9.3 to detect sex-biased dispersal, only the statistics F_{ST} and the associated statistic of relatedness (r) showed significant differences between sexes and indicated male-biased dispersal; however, differences in F_{ST} and r values were not very large between the sexes (Table 4.1). Although not significant, the assignment indices mAIc and vAIc showed a tendency for male-biased dispersal. However, F_{IS} was higher for females than males, showing a tendency for female-biased dispersal.

Table 4.1. Results from tests to detect sex-biased dispersal in red deer from the Scottish Highlands. Significance values for differences between males and females estimates were obtained after 10,000 permutations; tests with P < 0.05 are indicated in bold.

	n	F_{IS}	F_{ST}	Relatedness	mAIc	vAIc
Male	318	0.0271	0.0179	0.0342	-9.1636	98418.18
Female	250	0.0369	0.0247	0.0465	11.6561	1309.787
P-value		0.7635	**0.0045**	**0.0064**	0.07	0.21

Factorial correspondence analyses revealed a similar pattern of population structure for males and females with adjacent populations found to cluster closely in the plot (Fig. 4.2). For males the first three factors explained 48.77% of the observed genetic variance. Factor 1 clearly separated estates sampled at either side of the Great Glen for and explained 24.65 % of the genetic variance. Factor 2 explained 15.01% of the genetic variance and further separated populations of males sampled east to the Great Glen with the estates of CL and FL being distant to the rest of estates. Factor 3 explained 9.11% of the genetic variance. In females the three first factors explained 45.37% of the genetic variance. Factor 1 clearly separated the populations according to the side of the Great Glen that they were sampled and explained 21.24% of the genetic variance. Factor 2 explained 13.97% of the genetic variance and also indicated the differentiation of the estates of FL and CL from the rest of the estates sampled east of the Great Glen. Factor 3 explained 10.16% of the genetic variance.



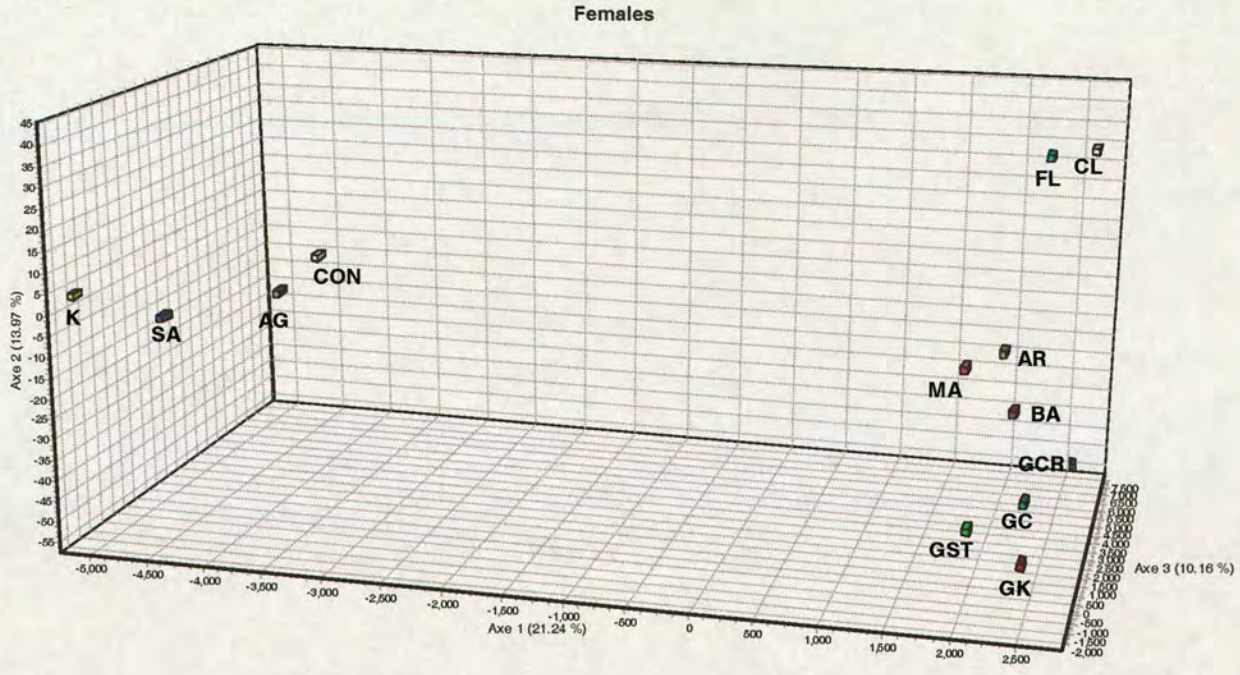


Figure 4.2. Three-dimensional plot illustrating results from the factorial correspondence analysis for male (previous page) and female (this page) genotypes. Colours and abbreviations different sampling areas (see Fig. 4.1 for sampling area codes).

4.5.2 Sex-biased dispersal in relation to spatial distance

Spatial autocorrelation analysis comparing kinship coefficients and geographical distances also indicated similar patterns for males and females, with an almost linear relationship with geographical distance (Fig. 4.3). The global regression slope between Fij and distance (0-112 km) was negative and significant both for males and females ($slope_m = -1.75x10^{-7}$, $P < 0.001$; $slope_f = -1.95x10^{-7}$, $P < 0.001$) (Figure 4.3). The proportion of the kinship coefficient variance explained by distance was very small ($r^2_m = 2.71\%$ and $r^2_f = 3.28\%$), as expected for the high sampling variance of the pair-wise kinship coefficients (Ritland 1996; Vekemans & Hardy 2004). For analyses conducted at different distance intervals, maximum kinship coefficients were obtained at the first distance interval (0-5 km) for both sexes, with average values of 0.01 and 0.012 for males and females, respectively (Figure 4.3).

Kinship coefficients were significantly positive for the first 3 intervals (up to 20 km) for females and the first 4 intervals (up to 30 km) for males. A significant negative association was obtained for both sexes for the 3 last intervals (from 60 to up over 100 km). Intermediate interval distances between 20-50 km for females and 40-70 km were close or within the 95% interval envelope of no association between genetic and spatial distances.

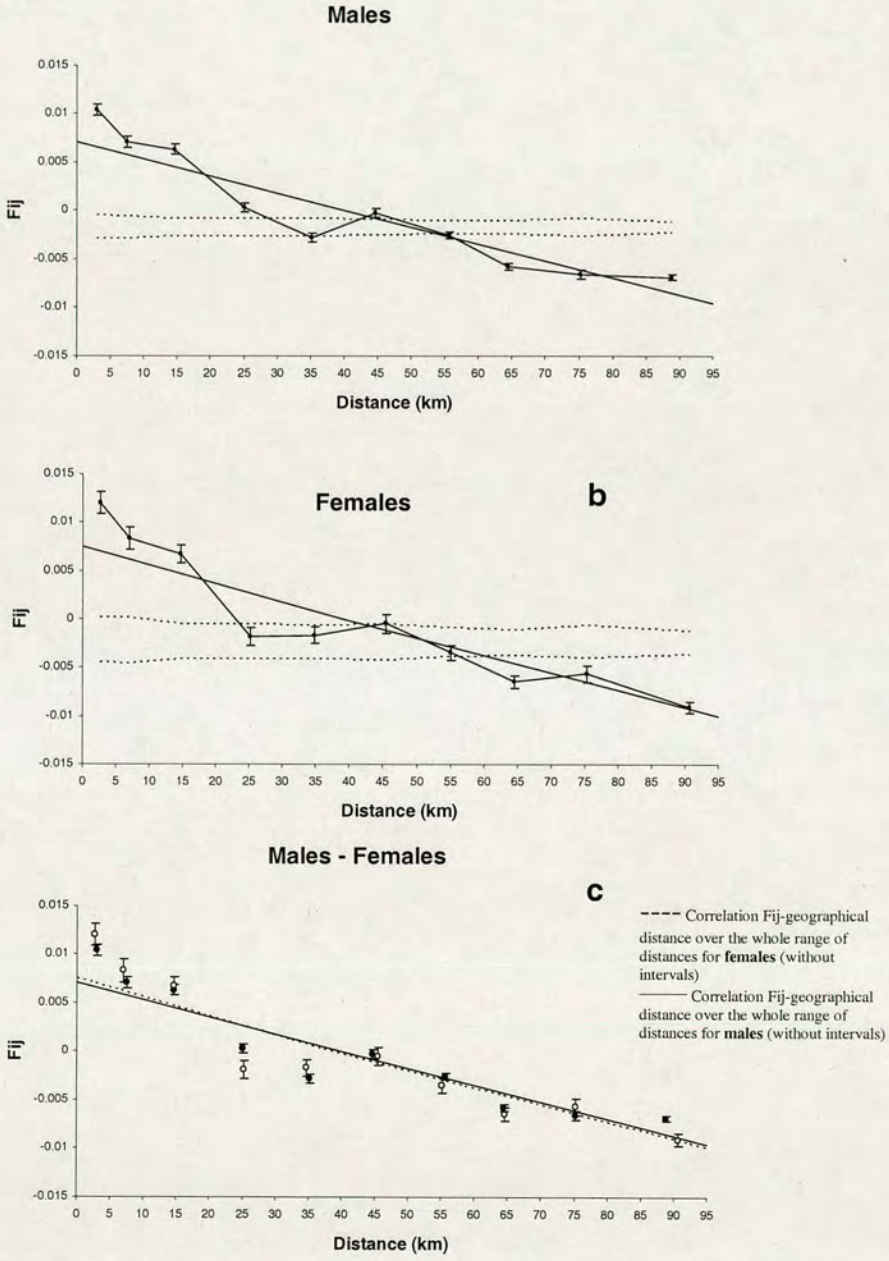


Figure 4.3. Correlograms showing the spatial autocorrelation for average kinship estimates for males (a), females (b), and males and females (c) at the different geographical distance intervals. Open circles represent female data and close circles represent male data. Standard deviations from the mean are indicated by error bars. The straight line in plot a & b illustrates the correlation between kinship and geographical distance over the whole range of geographical distances (without intervals). The 95% null hypothesis confidence intervals are indicated by the dotted lines in plot a & b.

Finer-scale spatial autocorrelation analysis only including individuals separated by 0-5 km revealed a very similar pattern for males and females (Fig. 4.4). Fij estimates were slightly higher for females at the first two intervals (up to 1km) and at the 3-4 km interval. Both sexes showed an oscillating pattern in Fij estimates, decreasing after 1 km and increasing and peaking again at distances of 3-4 km (Fig. 4.4).

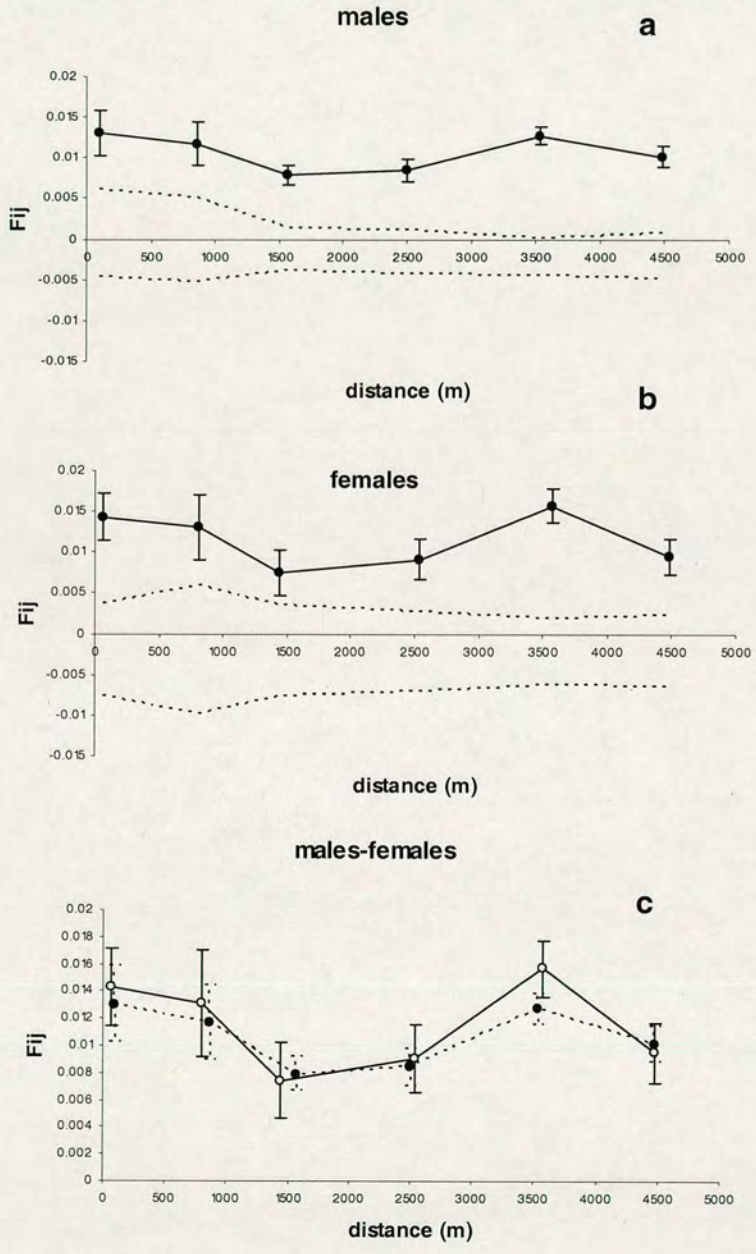


Figure 4.4. Correlograms showing the spatial autocorrelation for average kinship estimates at small distances (0-5 km). Standard deviations from the mean are indicated by error bars The 95% null hypothesis confidence intervals are indicated by the dotted lines. a) results for males, b) results for females and c) plot comparing both males and females. Key as for previous figure.

To assess the influence of sampling strategy on the analyses, the plots in figure 4.5 show to the distribution of pair-wise distances for males and females across the whole geographical distance range and for small distances. Comparisons of nearest neighbour distances significantly differed between males and females, with means of 738.85 m for males and 229 m for females ($t = 6.614$, $P < 0.001$).

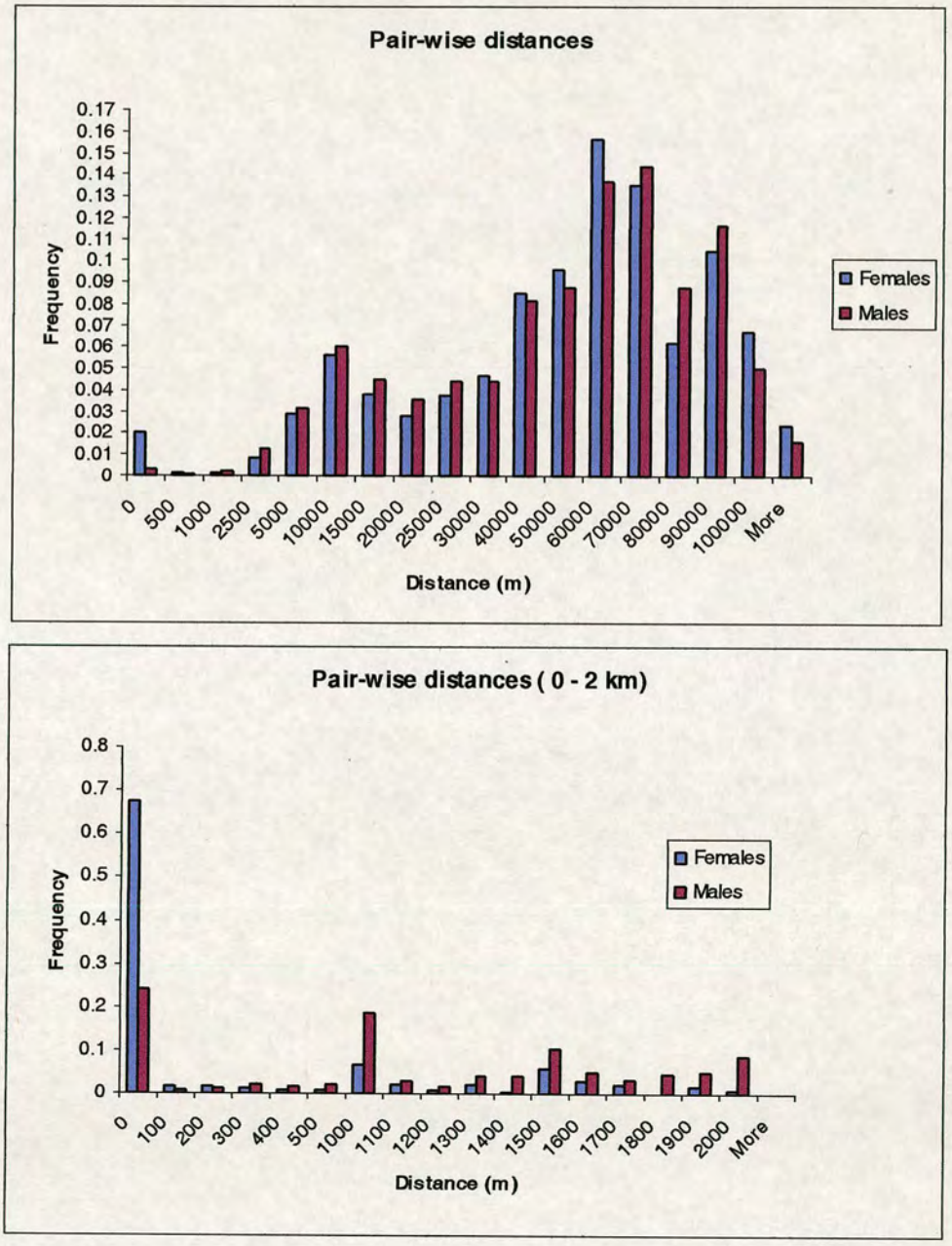


Figure 4.5. Distribution of pair-wise distances along the whole range of distances (top) and along distances 0-2 km (bottom).

4.6 Discussion

4.6.1 Detection of sex-biased dispersal

Despite strong male-biased dispersal is expected for mammal with polygynous mating system (Greenwood 1980; Dobson 1982), red deer dispersal in the Scottish Highlands was found to be only weakly male-biased. With the exception of F_{IS}, all other statistics used to detect sex-biased dispersal indicated a trend towards male-biased dispersal; however, only F_{ST} and relatedness showed significant differences between males and females but differences between sexes were not as extreme as expected. Simulation studies have shown that F_{ST} (and its associated relatedness measure) is the only statistic included in the tests powerful enough to detect a bias when dispersal rates are high and the bias is lower than 80:20 (Goudet *et al.* 2002), i.e. in the case of male-biased dispersal 80% of males disperse and only 20% of females disperse. Dispersal rates of red deer and thus gene flow in our relatively small study area are high as reflected by the relatively low F_{ST} values obtained (F_{STm} = 0.0179; F_{STf} = 0.0247). Although longer average dispersal distances for males than females have been recorded from recovery data from calf tagging programs in mainland Scotland (Daniels & McLean 2003), no detailed information exists about the rate of dispersal of each of the sexes. Dispersal rates recorded in the long-term study area of the North Block in Isle of Rum showed an average of 30% adult males and 17% adult females dispersing from their natal subdivision in the study area in a given year (Nussey *et al.* 2005), suggesting that the bias in dispersal rates might not be as high as originally thought for a strongly polygynous mammal. However, dispersal is a scale-dependent phenomenon and therefore dispersal on Rum might be limited by the size of the island and not comparable to that on the Scottish mainland (Clutton-Brock *et al.* 1989). Home ranges for red deer in Rum are smaller than those for red deer in the mainland (Clutton-Brock & Albon 1989) and due to the small size of the island of Rum dispersal distances will be much restricted than in the mainland (Clutton-Brock & Albon 1989). However, land fragmentation due to natural or man-made features found in mainland Scotland might make the movement of individuals difficult (see Chapter 3; Pérez-Espona *et al.* submitted).

Differences in the estimates of F_{ST} or relatedness found between sexes found in our study were much smaller than those found in other studies in which a strong sex-biased dispersal was indicated even at small scales (e.g. $F_{STmales}$= 0.15 and $F_{STfemales}$ = 0.03 in the European shrew *Crocidura russula* in a study in gardens of 2 km^2, Balloux *et al.* 1998; $relatedness_{males}$ = -0.225 and $relatness_{females}$ = 0.141 in the European rabbit in a study area of 1.5 ha, Surridge *et al.* 1999). Values for the other fixation index, F_{IS}, showed similar heterozygote deficiency with positive values for both sexes (only slightly higher for females) that could indicate a mix of resident and immigrant individuals were sampled for both male and female populations (Goudet *et al.* 2002).

Assignment tests (mAIc and vAIc) values did not significantly differ between sexes but both tests indicated a male-biased dispersal. Although assignment indices have proved very useful in detecting sex-biased dispersal in other species such as the European shrew (Favre *et al.* 1997) and bottlenose dolphins (*Tursiops aduncus*; Möller & Beheregaray 2004), other studies have also encountered difficulties in obtaining significant differences between sexes (e.g. white-footed mouse *Peromyscus leucopus*, Mossman & Waser 1999; coastal river otters *Lontra canadensis*, Blundell *et al.* 2002). These difficulties are associated with the fact that assignment tests work best at intermediate rates of dispersal (Goudet *et al.* 2002). If dispersal is common, immigrant and resident genotypes will be very similar and therefore it will be difficult to tell them a part; if dispersal is very rare then the probability of having sampled a migrant is very low. Simulation studies have shown that assignment tests cannot detect a bias when dispersal rates are higher than 10% (Goudet *et al.* 2002) and/or population differentiation is low ($F_{ST} < 0.05$) (Cornuet *et al.* 1999), which could explain the lack of significance in the assignment tests in this study.

Data on sex-biased dispersal of large polygynous mammals using bi-parentally inherited markers are almost absent, therefore comparisons of our results to other studies is difficult. In ungulates in particular, it is limited to a study of roe deer (*Capreolus capreolus*) in one population in south-western France in which female-

biased dispersal as predicted by theory for a resource-defence mating species was not found (Coulon *et al.* 2006). However, in this study it is not clear if comparisons between sexes were done including pre-dispersal individuals which could have obscured the signal of sex-biased dispersal.

The male-biased dispersal found in our study area using the fixation index statistics should be interpreted with caution as differences in the statistics between sexes were likely to be affected by the different sampling strategies adopted for males and females. Female samples were collected from fewer locations than were those of males and in some estates many females were collected in exactly the same location. These differences in the sampling strategies might have increased the chances of sampling more closely related females in comparison to males and might therefore have inflated some of the female F_{ST} and relatedness estimates between estates (see Fig. 4.1 & 4.5).

Factorial correspondence showed a similar pattern of spatial genetic population structure for both sexes. The same pattern, with major sub-structuring of the study area between estates sampled to either side of the Great Glen and the differentiation of FL and CL from those estates sampled east of the Great Glen, was also found in a previous study combining genotypes from both sexes of pre- and post-dispersal individuals (see chapter 3 or Pérez-Espona *et al.* submitted). This would indicate that sex-biased dispersal is not strong enough to influence population structure of red deer in the Scottish Highlands.

4.6.2 Sex-biased dispersal in relation to spatial distance

Spatial autocorrelation analyses revealed a similar pattern of spatial structure for both sexes in our study (Fig. 4.2). The distance at which the global regression line reached $F_{ij} = 0$ was about 40 km for males and females indicating no differences in the association between relatedness and distance between sexes. Furthermore, autocorrelation analyses revealed a similar significant positive association for both sexes between pair-wise kinship estimates and geographical distance for intervals

covering 0-20 km (Figure 4.2). This range of distance at which kinship estimates were positive agreed with some of the common dispersal distances reported in field-based studies (Daniels & McClean 2003; Sibbald *et al.* unpublished data). Overall mean dispersal distances recorded from calves tagged as neonates in four areas of mainland Scotland and shot or found dead as adults varied depending on the study area and ranged between 3.3-7.4 km for males and 1.9-3.5 km for females depending on the area. Maximal dispersal distances were similar for both sexes, 13.7-17.8 km for males and 12.3-16.7 km for females, and only two individuals, from the same estate, were found to disperse large distances, one male dispersing 57.6 km and one female 31 km (Daniels & McLean 2003). However, in the four areas studied some male and female individuals did not disperse from the natal area and, interestingly in one of the study areas (West Grampians) no difference in average dispersal distances was found between the sexes, both dispersing an average of 3.4 km (Daniels & McLean 2003). Although calf tagging data suggested that on average males disperse more than females, data from these calf tagging programmes did not differentiate between natal dispersal (permanent emigration of juveniles from the natal area) and temporary dispersal of mature males during the rut (Clutton-Brock & Albon 1989). A study in the Cairngorm area following 18 adult males and 4 females carrying GPS collars showed that, on average, males dispersed around 10 km from their normal over-wintering area (range 3-21 km) during the rutting period whereas hinds moved less than 3 km (Sibbald *et al.* unpublished data). Samples from males for our study were collected at the end of summer when adult males are supposed to have started to disperse to find female areas and thus the possibility of finding male-biased dispersal was maximised.

Kinship values did not differ strongly among sexes even at the small scale (0-5 km), even though nearest neighbour distances were significantly lower for females than for males. Within this short distance interval we were expecting to find the largest differences between sexes, as at short distances kin competition avoidance would be expected to be important (Lambin *et al.* 2001) and differences in spatial structure at shorter distances might be mainly due to differences in sex-specific dispersal rates and distances (Ronce *et al.* 2001). On the North Block study area on the island of

Rum where red deer are individually monitored, differences on genetic relatedness (measured with the Ritland & Linch 1999 coefficient) between male and female red deer were only found at extremely fine-scale (< 500 m), with female relatedness being high at very short distance intervals (< 100m) and steeply decreasing with distance, but with males relatedness not showing a correlation with geographical distance (Nussey *et al.* 2005). Although the scale of our study greatly differs from the one conducted in Rum, differences in relatedness between sexes in our study were not strong at the first interval (0-500 m) for which average inter-individual distances were 94.86 m for males and 66.45 m for females. Ecological, historical and anthropogenic factors affecting red deer biology on Rum and in the Scottish mainland are also important to have in mind in order to explain discrepancies in the findings. For example, current red deer in the isle of Rum are not native and have originated from several introductions including some from deer parks (Nussey *et al.* in press) and the North Block was released from culling in 1973 so individuals are more used to people in comparison to red deer in the mainland which are disturbed due to stalking and therefore are more prone to flee in human presence. Nonetheless, despite the fact that samples in our study were obtained from culling operations, if red deer male-biased dispersal was strong in the Scottish Highlands we would have expected much higher levels of relatedness between females than between males, in particular for distances < 500m.

Landscape features such as sea lochs, forests, roads and slopes have been shown to play an important role in red deer population sub-structuring in our study area (see chapter 3; Pérez-Espona *et al.* submitted). In particular, landscape features might have strongly influenced our spatial autocorrelation analysis for those distance intervals involving comparisons between individuals sampled at different sides of the Great Glen, where landscape features had the largest effect as gene flow barriers (see chapter 3; Pérez-Espona *et al.* submitted).

Competition between females at high densities can also affect the degree of spatial aggregation of related females and it has been shown to decrease in red deer in Rum (Albon *et al.* 1992; Nussey *et al.*2005), in white-tailed deer (*Odocoileus virginianus*)

in Canada (Kilpatrick *et al.* 2001) and in European wild rabbit (Surridge *et al.* 1999). In the Scottish Highlands red deer densities range from 5-25/ km^2 (Clutton-Brock & Albon 1989). If spatial aggregation of related females due to high density is low in our study area it could help explain the small kinship estimate difference between sexes at even the smallest distance interval. Furthermore, if competition for mates between males decreases as a consequence of high density of females and competition for food resources in both sexes increases due to population density increases, we could expect that the extent of male-biased dispersal could then potentially decrease (Perrin & Mazalov 2000).

The lack of strong male-biased dispersal found could also be explained if juvenile males disperse from their natal area to avoid competition for resources with females or competition for mates with older males, but then return to natal area or nearby areas to breed. This behaviour has been observed in the island of Rum, where males born in the North Block study area disperse and later on life return during the breeding season (Nussey *et al.* 2005). The majority of males in our study were sampled at the end of summer when males are expected to have dispersed to other estates in order to breed (rut or breeding dispersal) (Clutton-Brock & Albon 1989; Sibbald *et al.* unpublished data). Therefore, if some of the males included in our sampling were returning to the natal area to breed the signal of a natal male-biased dispersal would have been obscured. Subordinate males might also have difficulties in finding mates in new areas, so sometimes it could be less costly to stay near the natal area and wait there to find a mate (Johnson 1986).

Another male behaviour that could explain the weak male-biased dispersal observed is the fact that males generally move in groups (bachelor groups). If these bachelor groups are formed by some related individuals, kinship estimates within males and within females at a particular sampling location could potentially be similar. The behaviour of males dispersing with close relatives has, for example, been reported for lions (Packer *et al.* 1991).

4.7 Conclusions

Dispersal for red deer was found to be male-biased, as expected for a polygynous mammal (Greenwood 1980; Dobson 1982). However, the extent of male-biased dispersal was not very strong, even at short distances, indicating that although males might have higher dispersal rates or disperse longer distances than females, dispersal patterns are not very different between sexes and therefore sex-biased dispersal does not strongly influence red deer population structure in the Scottish Highlands. Several ecological, demographic and anthropogenic factors could be influencing the dispersal patterns of both male and female red deer in the Scottish Highlands. Therefore, future studies at a more local scale would help to assess to what extent each of these factors affect red deer sex-specific dispersal. The already observed behaviour in the Isle of Rum of males returning to the North Block to breed (Nussey *et al.* 2005) was also considered as a plausible cause for the relatively low degree of male-biased dispersal detected in this study. Re-examination of data from previous mark-recapture studies and new field studies in which male calves are tagged and followed into adulthood would provide information on the extent to which this male behaviour exists in mainland Scotland.

4.8 Acknowledgements

Red deer stalkers and managers from the estates of Ardgour, Ardverikie, Ben Alder, Clunes, Conaglen, Forest Lodge, Glencoe, Glencreran, Glenkinglass, Glenstrae, Kintail, Mamore and South Glen Affric are greatly thanked for the collection of samples. Russell Hooper is thanked for map reproduction and the calculation of inter-individual distances. The Macaulay Development Trust funded this project and supported S.P.-E. with a Ph.D. studentship. SEERAD (Scottish Executive) supported F.J. P.-B.

4.9 References

Albon SD, Staines HJ, Guiness FE, Clutton-Brock TH (1992) Density-dependent changes in the spacing behavior of female kin in red deer. *Journal of Animal Ecology* **61**, 131-137.

Balloux F, Goudet J, Perrin N (1998) Breeding system and genetic variance in the monogamous, semi-social shrew, *Crocidura russula*. *Evolution* **52**, 1230-1235.

Belkhir K. et al. GENETIX, logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Beyer, H.L. (2004) Hawth's Analysis Tools for ARCGIS available at http://www.spatialecology.com/htools

Blundell GM, Ben-David M, Groves P, Bowyers RT, Geffen E (2002) Characteristics of sex-biased dispersal and gene flow in coastal river otters: implications for natural recolonization of extirpated populations. *Molecular Ecology* **11**, 289-303.

Boudjemadi K, Lecomte J, Clobert J (1999) Influence of connectivity on demography and dispersal in two contrasting habitats: an experimental approach. *Journal of Animal Ecology* **68**, 1207-1224.

Cathey JC, Bickham JW, Patton JC (1998) Introgressive hybridization and nonconcordant evolutionary history of maternal and paternal lineages in North American deer. *Evolution* **52**, 1224-1229.

Chepko-Sade BD, Halpin ZT (1987) *Mammalian Dispersal Patterns: the Effects of Social Structure on Population Genetics* The University of Chicago Press, Chicago.

Christian SF (1994) Dispersal and other inter-group movements in badgers *Meles meles*. *Zeitschrift fuer Saeugetierkunde* **59**, 218-223.

Clutton-Brock TH, Guinness FE, Albon SD (1982a) *Behaviour and Ecology of Two Sexes* University of Chicago Press, Chicago.

Clutton-Brock TH, Albon SD, Guinness FE (1982b) Competition between female relatives in a matrilocal mammal. *Nature* **300**, 178-180.

Clutton-Brock TH, Albon SD (1989) *Red deer in the Highlands* BSP Professional Books, Oxford.

Clutton-Brock TH, Albon SD (1992) Trial and error in the Highlands. *Nature* **358**, 11-12.

Clutton-Brock TH, Coulson TN, Milner AD, Thomson D, Armstrong HM (2002) Sex differences in emigration and mortality affect optimal management of deer populations. *Nature* **415**, 633-637.

Clutton-Brock TH, Coulson T, Milner JM (2004) Red deer stocks in the Highlands of Scotland. *Nature* **429**, 261-262.

Cornuet J, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989-2000.

Coulon A, Cosson JF, Morellet N, *et al.* (2006) Dispersal is not female biased in a resource-defence mating ungulate, the European roe deer. *Proceedings of the Royal Society of London, Biological Series B* **273**, 341-348.

Coulson T, Albon SD, Guinness FE, Pemberton JM, Clutton-Brock TH (1997) Population substructure, local density, and calf winter survival in red deer (*Cervus elaphus*). *Ecology* **78**, 852-863.

Daniels M, McClean C (2003) Red deer calf tagging programmes in Scotland - an analysis. *Deer (The Journal of the British Deer Society)* **12**, 420-423.

Deer Commission for Scotland website. http://www.dcs.gov.uk

Dobson FS (1982) Competition for mates and predominant juvenile dispersal in mammals. *Animal Behaviour* **30**, 1183-1192.

Favre L, Balloux F, Goudet J, Perrin N (1997) Female-biased dispersal in the monogamous mammal Crocidura russula: evidence from field data and microsatellite patterns. *Proceedings of the Royal Society of London, Biological Series B* **264**, 127-132.

Gandon S, Michalakis Y (2001) Multiple causes of the evolution of dispersal In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD). Oxford University Press, Oxford.

Goudet J (1995) FSTAT version 1.2: a computer program to calculate F-statistics. *Journal of Heredity* **86**, 485-486.

Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited markers. *Molecular Ecology* **11**, 1103-1114.

Greenwood PJ (1980) Mating system, philopatry, and dispersal in birds and mammals. *Animal Behaviour* **28**, 1140-1162.

Hardy O, Vekemans X (2002) SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* **2**, 618-620.

Hartl DL, Clark AG (1997) *Principles of Population Genetics* Sinauer Associates, Sunderland, Massachussets.

Johnson CN (1986) Sex-biased philopatry and dispersal in mammals. *Oecologia* **69**, 626-627.

Johnson ML, Gaines MS (1990) Evolution of dispersal: Theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics* **21**, 449-480.

Kilpatrick HJ, Spohr SM, Lima KK (2001) Effects of population reduction on home ranges of female white-tailed deer at high densities. *Canadian Journal of Zoology* **79**, 949-954.

Koening WD, Van Vuren D, Hooge PN (1996) Detectability, philopatry, and the distribution of dispersal distances in vertebrates. *Trends in Ecology and Evolution* **11**, 514-517.

Kruuk H (1989) *The Social Badger: Ecology and Behavior of a Group Living Carnivore (Meles meles)* Oxford University Press, Oxford.

Lambin X, Aars J, Piertney SB (2001) Dispersal, intraspecific competition, kin competition and kin facilitation: a review of empirical evidence. In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD). Oxford University Press, Oxford.

Maehr DS, Land ED, Shindle DB, Bass OL, Hoctor TS (2002) Florida panther dispersal and conservation. *Biological Conservation* **106**, 187-197.

Möller LM, Beheregaray LB (2004) Genetic evidence for sex-biased dispersal. *Molecular Ecology* **8**, 1063-1067.

Mossman CA, Waser PM (1999) Genetic detection of sex-biased dispersal. *Molecular Ecology* **8**, 1063-1067.
Mysterud A, Pérez-Barbería FJ, Gordon IJ (2001) The effect of season, sex and feeding style on home range versus body mass scaling in temperate ruminants. *Oecologia* **127**, 30-39.
Nussey DH, Coltman DW, Coulson T, *et al.* (2005) Rapidly declining fine-scale spatial genetic structure in female red deer. *Molecular Ecology* **14**, 3395-3405.
Nussey DH, Pemberton J, Donald A, Kruuk LEB (in press) Genetic consequences of human management in an introduced island population of red deer (*Cervus elaphus*). *Heredity*.
Paetkau D (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* **4**, 347-354.
Pemberton JM, Albon SD, Guinness FE, Clutton-Brock TH, Dover GA (1992) Behavioural estimates of male mating success tested by DNA fingerprinting in a polygynous mammal. *Behavioral Ecology* **3**, 66-75.
Perrin N, Goudet J (2001) Inbreeding, kinship, and the evolution of natal dispersal. In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), p. 123-142. Oxford University Press., Oxford, UK.
Perrin N, Mazalov V (2000) Local competition, inbreeding, and the evolution of sex-biased dispersal. *The American Naturalist* **155**, 116-127.
Petit E, Mayer F (1999) Male dispersal in the noctule bat (*Nyctalys noctula*): where are the limits? *Proceedings of the Royal Society of London, Biological Series B* **266**, 1717-1722.
Pope LC, Doming-Roura X, Erven K, Burke T (2006) Isolation by distance and gene flow of the Eurasian badger (*Meles meles*) at both local and broad scale. *Molecular Ecology* **15**, 371-386.
Pope TR (1999.) Effects of demographic change on group kin structure and gene dynamics of populations of red howling monkeys. *Journal of Mammalogy*, 692–712.
Prugnolle F, De Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity* **88**, 161-165.
Pusey AE, Packer C (1987) Dispersal and philopatry. In: *Primate Societies* (eds. B. B. Smuts, Cheney DL, R. M. Seyfarth, Wrangham RW, Struhsaker TT), pp. 250-266. University of Chicago Press, Chicago.
Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution* **43**, 258-275.
Red Deer Commission (1978) Annual Report for 1977, Edinburgh, HMSO.
Red Deer Commission (1982) Annual report for 1981, Edinburgh, HMSO.
Red Deer Commission (1983) Annual report for 1982, Edinburgh, HMSO.
Red Deer Commission (1989) Annual report for 1988, Edinburgh, HMSO.
Richardson BJ, Hayes RA, Wheeler SH, Yardin MR (2002) Social structures, genetic structures and dispersal strategies in Australian rabbit (Oryctolagus cuniculus) populations. *Behavioral Ecology and Sociobiology* **51**, 113-121.
Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetical Research* **67**, 175-185.

Ronce O, Olivieri I, Clobert J, Danchin E (2001) Perspectives on the study of dispersal evolution In: Dispersal (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), p. 341–357. Oxford University Press, Oxford.

Rogers LM, Delahay R, Cheeseman CL (1998) Movement of badgers (*Meles meles*) in a high-density population: individual population and disease effects. *Proceedings of the Royal Society of London, Series B, Biological Sciences* **265**, 1269-1276.

Roper TJ, Ostler JR, Conradt L (2003) The process of dispersal in badgers *Meles meles*. *Mammal Review* **33**, 314-318.

Rubin ES, Bleich VC (2005) Sexual segregation: a necessary consideration in wildlife conservation. In: Sexual segregation in vertebrates: ecology of the two sexes (eds. Ruckstuhl KE, Neuhaus P), p. 379-391. Cambridge University Press, Cambridge.

Seielstad MT, Minch E, Cavalli-Sforza L (1998) Genetic evidence for a higher female migration rate in humans. *Nature Genetics* **20**, 278-280.

Shaw CN, Wilson PJ, White BN (2003) A reliable molecular method of gender determination for mammals. *Journal of Mammalogy* **84**, 123-128.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology & Systematics* **16**, 393-430.

Slatkin M (1987) Gene flow and the geographical structure of natural populations. *Science* **236**, 787-792.

Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* **82**, 561-573.

Sokal RR, Wartenberg D (1983) A test of spatial autocorrelation using an isolation-by-distance model. *Genetics* **105**, 219-237.

Stenseth NC, Lidicker WZ (1992) The study of dispersal: a conceptual guide. In: *Animal dispersal: small mammals as a model* (eds. Stenseth NC, Lidicker WZ), pp. 5-20. Chapman and Hall, New York, USA.

Surridge AK, Bell DJ, Hewitt GM (1999) From population structure to individual behaviour: genetic analysis of social structure in the European wild rabbit (*Oryctolagus cuniculus*). *Biological Journal of the Linnean Society* **68**, 57-71.

Tuyttens FAM, Delahay RJ, MacDonald DW, *et al.* (1995) Spatial perturbation caused by a badger (*Meles meles*) culling operation: implications for the function of territoriality and the control of bovine tuberculosis (*Mycobacterium bovis*). *Journal of Animal Ecology* **69**, 815-828.

Tuyttens FAM, Macdonald DW (2000) Consequences of social perturbation for wildlife management and conservation. In: *Behaviour and conservation* (eds. L. M. Gosling, Sutherland WJ), pp. 315-329. Cambridge University Press, Cambridge.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* **13**, 921–935.

Wahlstrom LK, Liberg O (1994a) Patterns of dispersal and seasonal migration in roe deer (*Capreolus capreolus*). *Journal of Zoology* **235**, 455-467.

Waser PM, Creel SR, Lucas JR (1994) Death and disappearance: estimating mortality risks associated with philopatry and dispersal. *Behavioral Ecology and Sociobiology* **5**, 135-141.

Whitlock MC (2001) Dispersal and the genetic properties of metapopulations. In: Dispersal (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), p. 273–282. Oxford University Press, Oxford.

Wiens JA (2001) The landscape context of dispersal. In: Dispersal (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD), p. 96-109. Oxford University Press, Oxford

Wolff JO (1993) What is the role of adults in mammalian juvenile dispersal? *Oikos* **68**, 173-176.

Woodroffe R, MacDonald D, Silva JD (1995) Dispersal and philopatry in the European badger, *Meles meles*. *Journal of Zoology* **237**, 227-239.

Zenger KR, Eldridge MDB, Cooper DW (2003) Intraspecific variation, sex-biased dispersal and phylogeography of the eastern grey kangaroo (Macropus giganteus). *Heredity* **91**, 153-162.

Chapter 5

Genetic diversity, population structure and sex-biased dispersal of Scottish red deer (*Cervus elaphus*) inferred by mitochondrial DNA control region sequences

S.Pérez-Espona,*§ F.J. Pérez-Barbería, § W. Goodall-Copestake, C. Jiggins, * I.J. Gordon, † J.M. Pemberton*

*Institute of Evolutionary Biology, School of Biological Sciences,The University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK

§ The Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK

† Sustainable Ecosystems, CSIRO – Davies Laboratory, PMB PO Aitkenvale, Qld 4814, Australia

Corresponding author:

Sílvia Pérez-Espona,

Institute of Evolutionary Biology, School of Biological Sciences, The University of Edinburgh

The King's Buildings, West Mains Road, Edinburgh EH9 3J, UK

Fax: ++ 44 (0)131 650 6564

Tel: ++44 (0)131 651 3612

E-mail: s.perez-espona@sms.ed.ac.uk

Author's contribution:

SPE organised the sampling, conducted the laboratory work and performed the statistical analyses. WGC provided assistance in the sequence alignment and phylogenetic analyses. The project was set up by FJPB and IJG. The whole project was supervised by JPB, JMP, CDJ and IJG. All authors critically read and improved drafts of this manuscript.

5.1 Abstract

Scottish Highland red deer (*Cervus elaphus*) represent the largest population of red deer in Europe. However, little is known about their genetic diversity and population structure. In this study the genetic diversity and population structure of red deer in the Scottish Highlands is assessed by sequencing ≥ 821bp of the mitochondrial control region (mtDNA CR) of 695 individuals previously genotyped for 21 microsatellite markers. Genetic diversity was high, with 74 haplotypes found in the study area. Phylogenetic analyses of the 74 haplotypes and 47 sequences of other red deer and species within the Cervinae downloaded from Genbank revealed that no individuals sampled in this study had mtDNA haplotypes from foreign species or subspecies of deer, but suggested that there may have been a few localised red deer translocations among British localities. Significant mtDNA CR population structure was found in the study area ($\Phi_{ST} = 0.3483$), with 18.25% of the genetic variation explained by differences between the four regions in which the study area was divided. A minimum spanning network and inferences of historical demographic processes suggested a relatively recent population expansion for Scottish Highland red deer that may reflect possible colonisation events of mainland Scotland during the Pleistocene. Patterns of population differentiation were similar to those found in a previous study using microsatellite markers, with nearby populations not being significantly differentiated from each other. Comparisons of population structure estimated from mtDNA CR sequences and microsatellite markers, and comparisons of mtDNA CR population structure estimates between the sexes indicated that male-biased dispersal is weak across the study area and occurs at a small scale (between nearby populations). Effective longer distance movements may be predominantly due to females.

Keywords: *Cervus elaphus*, mitochondrial DNA, control region, red deer, introgression, sex-biased dispersal, population expansion, population structure.

5.2 Introduction

The red deer (*Cervus elaphus*) is one of the most widespread and best studied deer species in the world (Clutton-Brock & Albon 1989; Ludt *et al.* 2004). In Europe, red deer populations have been strongly influenced by human management over centuries, either indirectly through habitat destruction and grazing competition with livestock or directly through selective hunting and translocations of foreign stocks in order to improve trophy hunting (Lowe & Gardiner 1974; Hartl *et al.* 2003; Feulner *et al.* 2004). The largest population of red deer in Europe occurs in the British Isles (30 % of the total European population), with the majority of the population found in Scotland (Clutton-Brock *et al.* 1989).

Red deer are known to have occurred in Scotland throughout the mid-interglacial temperate woodland phases of the Middle-Upper Pleistocene (c.730,000-130,000 years BP) with a continuous presence since the end of the last glaciation (c.11,000 years BP) (Lister 1984) when forests interspersed with open areas were abundant throughout Scotland (Clutton-Brock & Albon 1989). However, due to over-hunting and deforestation in the Lowlands of Scotland associated with the development of farming cultures (c. 5,000 years BP), red deer were gradually displaced northwards to the mountainous Highlands and were almost absent in the Lowlands (and in much of England and Wales) by the end of the Middle Ages (Whitehead 1964; Lister 1984). Continuous deforestation and over-hunting during the 16th-18th Centuries is thought to have caused a large decline in Scottish red deer populations with numbers being at the lowest by the second half of the 18th Century when large numbers of sheep were brought into the Highlands, strongly competing with red deer for grazing (Clutton-Brock & Albon 1989). The range and abundance of red deer however, rose again in the 19th Century due to a growing interest in deer hunting coupled with a decrease in profits from sheep rearing, which allowed large areas of land to be inhabited by deer (Lowe & Gardiner 1974; Clutton-Brock & Albon 1989). Although translocations of red deer among European countries have probably occurred since Roman times (Long 2003), it was during the 19th century that numerous translocations into Scotland of red deer stocks, in particular from English deer parks,

took place in order to improve the body and antler size for trophy hunting (Whitehead 1960, 1964; Lowe & Gardiner 1974). Furthermore, hybridisation of red deer with exotic species such as the North American wapiti (*Cervus elaphus canadensis*) and the Japanese sika deer (*Cervus nippon*) have also been documented (Whitehead 1960, 1964; Abernethy 1994; Goodman *et al.* 1999) threatening the native gene pool of Scottish red deer. Although red deer numbers probably also declined again due to hunting and poaching during periods of war (1939-1945), concerns about the protection of red deer coupled with a series of mild winters and lower competition with sheep helped to increase red deer numbers again, with the population thought to have doubled in the past 40-50 years (Clutton-Brock & Albon 1989; Clutton-Brock & Albon 1992; Clutton-Brock *et al.* 2004). Red deer are distributed across the Highlands, Islands and SW of Scotland and the highest densities are found in the Highlands where about 400,000 individuals are thought to occupy an area of 300,000 km^2 of the Scottish mainland (Clutton-Brock & Albon 1989; Clutton-Brock *et al.* 2004). Currently, the Deer Commission for Scotland oversees the management of deer in Scotland which mostly takes place by traditional stalking of stags, which continues to be an important source of revenue in the Scottish Highlands, and culling of hinds for population regulation (Deer Commission for Scotland website – http://www.dcs.gov.uk).

The use of molecular markers to assess the genetic diversity and population structure of a species provides important insights into its evolution and ecology, information which is valuable in order to develop science-based management or conservation policies (Moritz 1994; Moritz 2002; Crandall *et al.* 2000). For game species such as red deer, that have long been influenced by man, molecular markers have also helped to identify native stocks and evaluate the extent to which genetic diversity and population genetic structure has been affected by different human activities and past management practices (e.g. Hartl *et al.* 2003; Hartl *et al.* 2005; Zachos *et al.* 2003; Hmwe *et al.* in press; Nussey *et al.* in press).

Despite Scottish Highland red deer representing the largest population of red deer in Europe, studies assessing the genetic diversity and geographical structure of this

population are limited. A recent study using 21 microsatellite markers to assess gene flow of 695 Scottish Highland red deer showed high genetic diversity and significant population structure in the study area, with landscape features playing an important role in population differentiation (see Chapter 3 or Pérez-Espona *et al.* submitted).

For studies from which useful guidelines for management or conservation can be derived, it is important to acquire information from different genetic markers covering different evolutionary time scales (Crandall *et al.* 2000). Mitochondrial DNA markers, such as the control region, have been widely used for population studies as they are a powerful tool to infer population history (Avise *et al.* 1994) as it gives a more historical perspective on the evolution and demography of population than microsatellite markers (Paetkau *et al.* 1987). Phylogenetic relationships and the geographical distribution of mitochondrial DNA haplotypes can also be used to infer the extent to which human activities have affected the native gene pools and the population structure of a species (e.g. Eizirik *et al.* 2001; Kasapidis *et al.* 2005; Pereira *et al.* 2005; Thulin *et al.* 2006). Furthermore, due to their maternal inheritance in mammals, mtDNA markers reflect female dispersal and, therefore, comparisons of patterns of structure obtained with mtDNA and bi-parentally inherited markers such as microsatellites can provide information on sex-specific dispersal patterns (e.g. Ennos 1994; FitzSimmons *et al.* 1997; Escorza- Treviño & Dizon 2000; Piertney *et al.* 2000; reviewed in Prugnolle & De Meeus 2002).

On the Island of Rum (Scotland; Fig. 5.1), where long-term research on several aspects of red deer biology has been conducted, mtDNA haplotype diversity was found to be low, and some individuals presented a divergent haplotype related to Corsican red deer (*Cervus elaphus corsicanus*) both features presumably as a result of the entire population having originated from documented introduction events (Nussey *et al.* in press). MtDNA diversity and structure involving red deer from the mainland of Scotland, in particular in the Highlands where a more natural population exists, has been only assessed recently (Hmwe *et al.* in press). In a study involving 69 individuals from six Scottish and one English red deer populations, Hmwe *et al.* (in press) found a lack of congruence between geographical and genetic structure that

they attributed to past red deer management practices in Britain. However, in their study, evidence for the influence of human management on populations of red deer was only clear on two islands (Arran and Islay) (Fig. 5.1).

In this study, sequences of the mtDNA control region of the same 695 red deer individuals previously genotyped for 21 microsatellite were analysed to assess (i) the extent of mtDNA introgression of foreign deer subspecies or species and translocations of red deer stocks in the Scottish mainland, (ii) the genetic diversity and population structure of mtDNA in the Scottish mainland, (iii) past population demographic processes such as bottlenecks or population expansions and, (iv) the extent of sex-biased dispersal through comparisons mtDNA control region population structure with previous estimates of population structure obtained with microsatellite markers and through comparisons of mtDNA population structure estimates among the sexes.

5.3 Material and Methods

5.3.1 Study area and sample collection

The study area comprises a series of estates distributed across an east-west transect in the Scottish Highlands (Fig. 5.1). Samples consisted of an ear tip or a sample of jaw muscle from a total of 695 red deer (400 males and 295 females) legally shot on 14 open hill estates during the 2003-2004 hunting (culling) season (Fig. 5.1). Tissue samples were stored either in a -20ºC freezer or in tubes containing 100% ethanol.

5.3.2 DNA laboratory procedures

5.3.2.1 DNA extraction and sex determination

Genomic DNA was extracted from ear or jaw muscle using the DNAace Spin Tissue Mini Kit (Bioline) or with the DNEasy Tissue Kit™ (QIAGEN), following the

manufacturer's instructions. Sex determination of the samples was conducted by using a genetic sex marker (Shaw *et al.* 2003; see Chapter 4).

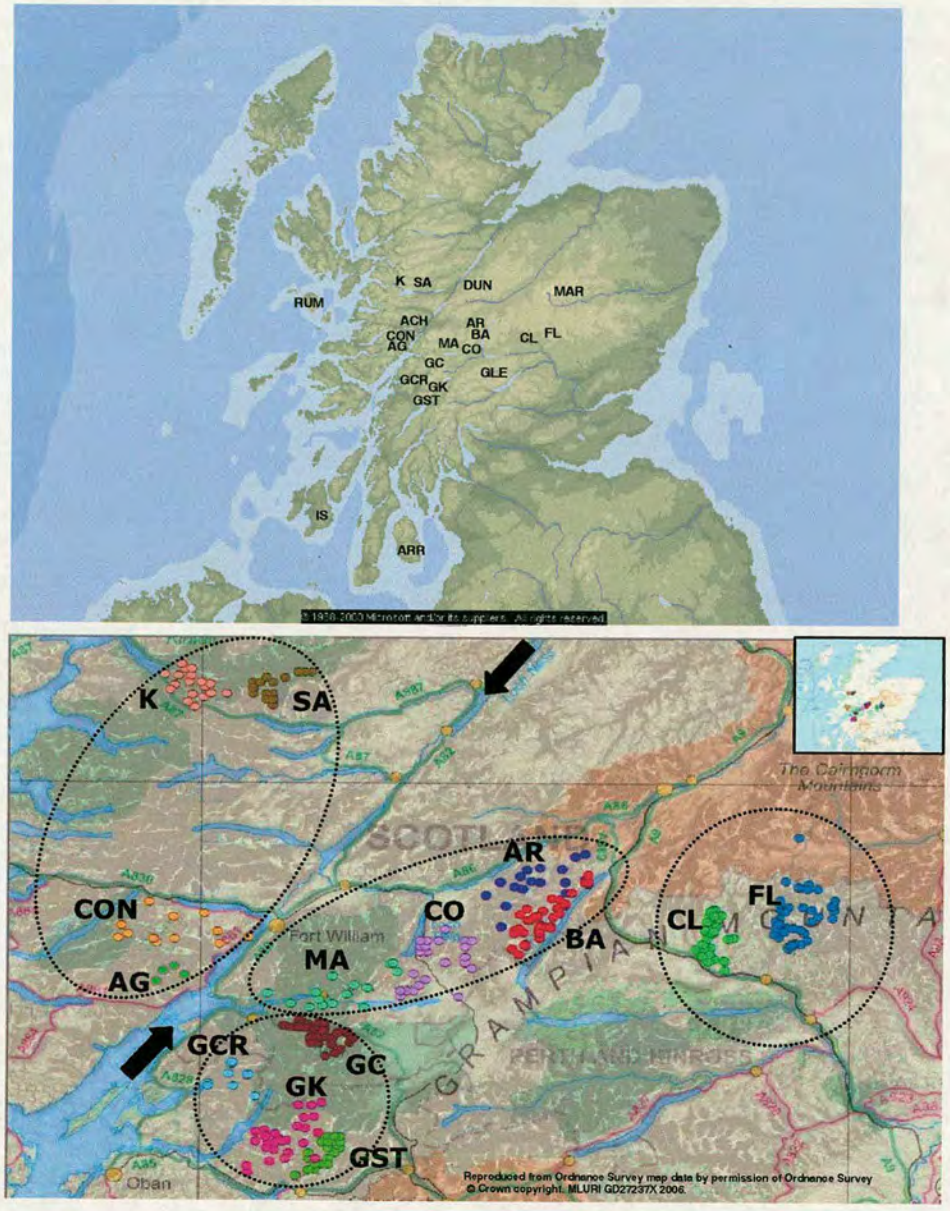


Fig. 1 Top Map showing the locations of populations included in this study and locations of populations included in the other mitochondrial DNA surveys of Scottish red deer (Hmwe *et al.* in press; Nussey *et al.* in press). **ACH** = Achnacarry, **ARR** = Arran, **DUN** = Dunachton, **GL**= Glen Lyon, **IS** = Islay, **MAR** = Mar, **RUM** = Rum. **Bottom** Map of the sampling area indicating the four main regions identified by microsatellite data (dashed line). Dots in the map do not correspond to the total of individuals sampled but to the culling locations (several individuals were sampled in the same location). Black arrows indicate location of the Great Glen. Dashed lines indicate the four main regions found in a previous study using microsatellite markers. **FL** = Forest Lodge, **CL**= Clunes, **BA**= Ben Alder, **AR**= Ardverikie, **CO**= Corrour, **MA**= Mamore, **GC**= Glencoe, **GCR**= Glencreran, **GK**= Glenkinglass, **GST**= Glenstrae, **CON**= Conaglen, **AG**= Ardgour, **SA**= South Glen Affric, **K**= Kintail.

5.3.2.2 Sequencing of the mitochondrial control region

The mitochondrial control region (mtDNA CR) and the partial flanking region of tRNA genes were amplified using the primers CST2 and CST39 (Polziehn *et al.* 1998). Amplification of the mtDNA CR was conducted in 50μl PCR reactions using 10-15ng of DNA, 1xNH4 Buffer, 1.5mM $MgCl_2$, 0.6mM of each primer, 1 unit of BIOTAQ™ polymerase (Bioline) and double processed tissue culture distilled H_2O (Sigma) to bring the volume up to 50 μl. The PCR cycling protocol for the amplification of the mtDNA CR involved an initial activation step of 94°C for 3 min, a 3 step-cycling consisting of a denaturing step of 94°C for 30 sec, annealing at 56°C for 30sec, ramping at 0.3°C/sec to an extension step of 72°C of 1min. The cycle was repeated 29 times and was followed by a final extension of 72°C for 10 min. PCR products were run in a 1% agarose gel to check if amplification was successful. Successful amplifications were purified using the Sigma-Genosys Genelute ™ PCR Clean-up Kit following manufacturer's instructions. A fragment of approximately 1,000bp of the mtDNA CR was sequenced in two reactions using 4 μl of purified PCR product, 2μl of the reaction mix DYEnamic ET Terminator Cycle Sequence Kit (Amersham Science) and 3μl of each of the primers CST2 and CST29 (Polziehn & Strobeck 2002). Sequencing cycling consisted of 25 cycles including a denaturation step of 95°C for 20 sec, an annealing step of 50°C for 15 sec, and a extension step of 60°C for 1 min. Sequences were run on a capillary ABI 3730 DNA Analyzer sequencer (Applied Biosystems).

5.4 Data analysis

5.4.1 Phylogenetic analyses

Sequences were manually edited using SEQUENCE NAVIGATOR version 1 (Applied Biosystems) and aligned using the Clustal V method implemented in the program MEGALIGN version 5 (distributed by DNAstar) and manually modified. Haplotype assignments to individuals were made using the redundant haplotypes search in MACCLADE version 4.0 (Maddison & Maddison 2000).

In order to assess the influence of translocations and hybridisations of red deer with other species, phylogenetic analyses were conducted on the haplotypes found in this study and 47 other mitochondrial control region sequences of red deer and other species within the Cervinae subfamily downloaded from Genbank (details are listed in Table 5.1). Alignment of our haplotypes and the 47 sequences from Genbank was conducted with the web-based software MUSCLE (Edgar 2004). Gaps corresponding to inferred insertion and deletion events by the automatic alignment were manually altered in MEGALIGN v. 5 to maximise DNA character positional homology. Alignment of ambiguous regions, microsatellite repeats, and positions for which the majority of samples contained a gap were excluded from subsequent analyses.

Phylogenetic analyses were conducted using Bayesian inference and maximum parsimony. MRBAYES version 3.1 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003) was used for the Bayesian inference. Three simultaneous independent analyses were run, each starting from a random tree with the default option of one cold and three incrementally heated chains. The chains were run for 1,500,000 generations, sampled every 100 generations, with 500,000 generations discarded as a conservative burn-in. Inspection of the posterior parameter values using TRACER (Rambaut & Drummond 2004) indicated that stationarity was reached by 500,000 generations. Default values were left for the other parameters, except for the model of DNA evolution that was set to a General Time Reversible (GTR) model with gamma distributed rate heterogeneity (Rodríguez *et al.* 1990). Maximum likelihood ratio tests implemented in MODELTEST v. 3.06 (Posada & Crandall 1998) indicated that a GTR model with a proportion of invariant sites plus gamma distributed rate heterogeneity was the most appropriate model. However, this more complex model of DNA evolution was not used because it did not appear to reach stationarity in the posterior distribution. The alpha value of the gamma parameter and proportion of invariant sites parameter appeared to be antagonistic when they were compared using TRACER. A majority rule phylogram was generated from the 30,003 post-burnin samples from all three independent analyses.

Maximum parsimony tree searches were conducted using PHYLIP version 3.65 (Felsenstein 2005). Gaps were treated as missing data and the search options were left at default values (equal weights, thorough tree and 10,000 saved trees). To generate support values, invariant characters were removed and 1,000 bootstrap pseudo-replicates were generated saving 10 trees per replicate.

Table 5.1. Details for the sequences downloaded from Genbank.

Species name	Common Name	Origin	Genbank Accession no.	Reference
Cervus elaphus scoticus[1]	Red deer (Rum A)	Isle of Rum (Scotland)	DQ386106	Nussey *et al.* in press
C. elaphus scoticus	Red deer (Rum B1)	Isle of Rum (Scotland)	DQ386107	Nussey *et al.* in press
C. elaphus scoticus	Red deer (Rum B2)	Isle of Rum (Scotland)	DQ386108	Nussey *et al.* in press
C. elaphus scoticus	Red deer (Rum B3)	Isle of Rum (Scotland)	DQ386109	Nussey *et al.* in press
C. elaphus scoticus	Red deer (Rum B4)	Isle of Rum (Scotland)	DQ386110	Nussey *et al.* in press
C. elaphus scoticus	Red deer (H1)	Glenartney (Scotland)	DQ452075	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H2)	Glenartney/ Isle of Islay (Scotland)	DQ452076	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H3)	Euston (England)	DQ452077	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H4)	Euston (England)/Isle of Arran (Scotland)	DQ452078	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H5)	Mar (Scotland)	DQ452079	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H6)	Mar/Dunachton (Scotland)	DQ452080	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H7)	Mar (Scotland)	DQ452081	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H8)	Mar (Scotland)	DQ452082	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H9)	Achnacarry (Scotland)	DQ452083	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H10)	Achnacarry/Dunachton (Scotland)	DQ452084	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H11)	Achnacarry (Scotland)	DQ452085	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H12)	Isle of Arran (Scotland)	DQ452086	Hmwe *et al.* in press
C. elaphus scoticus	Red deer (H13)	Dunachton (Scotland)	DQ452087	Hmwe *et al.* in press
C. elaphus	Red deer 21	Unknown	AF296816	Polziehn & Strobeck 2002
C. elaphus	Red deer 11	Unknown	AF296815	Polziehn & Strobeck 2002
C. albirostris	Thorold's white-lipped deer	San Diego Zoo (USA)	AF016972	Polziehn & Strobeck 1998
C. albirostris	Thorold's white-lipped deer	San Diego Zoo (USA)	AF016973	Polziehn & Strobeck 1998
C. e. atlanticus	Norwegian red deer	Norway	AF291888	Randi *et al.* 2001
C. e. hispanicus	Spanish red deer	Spain	AF291889	Randi *et al.* 2001
C. e. hippelaphus	European red deer 031	South of Italy	AF291886	Randi *et al.* 2001
C. e. hippelaphus	European red deer 038	North of Italy	AF291887	Randi *et al.* 2001
C. e. corsicanus	Corsican red deer	Sardinia (Italy)	AF291885	Randi *et al.* 2001
C. e. barbarus	Barbary red deer	Algeria	AF296808	Polziehn & Strobeck 2002
C. e. bactrianus	Bactrian red deer	Uzbekistan	AF296823	Polziehn & Strobeck 2002
C .e. yarkandensis	Yarkand wapiti	Tarim	AB074554	Mahmut *et al.* 2002
C. e. canadensis	Eastern Wapiti	Unknown	AY970666	Lee unpublished
C. e. nelsonii	Rocky Mountain wapiti	Rocky Mountain Banff (Canada)	AF016964	Polziehn & Strobeck 1998
C. e. nannodes	Tule elk	California (?)	AF016976	Polziehn & Strobeck 1998
C. e. roosevelti	Roosevelt's wapiti	Oregon (USA)	AF016971	Polziehn & Strobeck 1998
C. e. manitobensis	Manitoban wapiti	Manitoban Riding Mountain (Canada)	AF016960	Polziehn & Strobeck 1998
C. e. sibericus	Siberian wapiti	Unknown	AF058371	Polziehn & Strobeck 1998
C. e. songaricus		Hongshan (China)	AB074545	Mahmut *et al.* 2002
C. e. xanthopygus	Manchurian wapiti		AF296817	Polziehn & Strobeck 2002
C. e. wallichi		Tibet	AB074553	Mahmut *et al.* 2002
C. e. kansuensis	Kansu red deer	China	AF296819	Polziehn & Strobeck 2002
C. e. alashanicus	Alashan wapiti	China	AF296818	Polziehn & Strobeck 2002
C. nippon	Sika deer (sika 74)	Unknown (Japan?)	AF16974	Polziehn *et al.* 1998
C. nippon	Sika deer (sika 75)	Unknown (Japan?)	AF16975	Polziehn *et al.* 1998
Elaphurus davidianus	Père David's deer	China (Beijing)	AF291894	Randi *et al.* 2001
Axis porcinus	Hog deer	Indonesia	AF291897	Randi *et al.* 2001
Cervus eldi thamin	Manipur thamin	Conservation & Research Center (USA)	AY137123	Balakrishnan *et al.* 2003
Dama mesopotamica	Persian fallow deer	Iran	AF291896	Randi *et al.* 2001

[1] Haplotype Rum A from the island of Rum was found to be a haplotype closely related to *C. e. corsicanus*

5.4.2 Population analyses

Estimates of genetic diversity, population structure and inference of past demographic events were conducted in ARLEQUIN version 2.0 (Schneider *et al.* 2000). Genetic diversity indices consisted of haplotype diversity (h), nucleotide diversity (π) and number of polymorphic sites (s). Relationships between haplotypes were assessed by constructing a haplotype minimum spanning network using the program TCS version 1.21 (Clement *et al.* 2000). Population structure across the study area and between all sampling sites was assessed using an analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992). Pair-wise population differentiation was calculated using Φ-analogues of Wright's F-statistics (Wright 1951). Significance values were obtained after 1,000 permutations among populations using Fisher's exact tests (Ryman & Jorde 2001), with a Bonferroni correction applied for multiple comparisons (Rice 1989). Hierarchical AMOVAS were conducted to assess the distribution of the genetic variation among regions (Φ_{CT}) and populations within regions (Φ_{SC}). Populations were grouped in four regions (East, West, Central 1, Central 2) according to previous findings of population subdivision using microsatellite markers (Fig. 5.1, see Chapter 3, Pérez-Espona *et al.* submitted). Statistical significance for AMOVA analyses was obtained after 1,000 permutations (Excoffier *et al.* 1992).

Mitochondrial DNA markers are thought to reflect female dispersal due to their maternal inheritance. However, the inclusion of male migrants in the sampling has been shown to have a small but statistically significant effect when estimating population genetic structure using mtDNA markers (Tiedemann *et al.* 2000). Weak male-biased dispersal at small scales of the study area was found in a previous study comparing microsatellite data between post-dispersal males and females in a previous study (see Chapter 3 or Pérez-Espona *et al.* submitted). Therefore, population structure measures were also calculated for each sex separately in order to assess the influence of male dispersal on the distribution of mtDNA haplotypes.

Inferences regarding past demographic effects on the genetic variation in current Scottish Highland red deer were conducted by considering the whole study area as a

single population. Mismatch distributions of pair-wise nucleotide differences between haplotypes were compared with those expected under a sudden population expansion model (Slatkin & Hudson 1991; Rogers & Harpending 1992; Rogers 1995). Unimodal distributions are expected for populations that recently expanded or experienced a bottleneck, as individuals within a population will present similar haplotype divergence (in terms of nucleotide differences) (Slatkin & Hudson 1991; Rogers & Harpending 1992). In contrast, a multimodal or 'ragged' distribution is expected for a stable or slowly declining population (Slatkin & Hudson 1991). Statistical significance for the mismatch distributions was obtained using a goodness-of-fit test based on the sum of squared deviations between the observed and expected distributions (Schneider & Excoffier 1999) and the Harpending's raggedness index *rg* (Harpending 1994) after 1,000 simulations using the estimated parameters of the expected distribution for a population expansion. Population expansions were further assessed using the neutrality test *Fs* (Fu 1997) which has been shown to be more powerful than mismatch distributions in detecting population growth (Ramos-Onsins & Rozas 2002). For populations that have had stable size (in equilibrium), *Fs* is expected to be 0. Large significant and negative values of *Fs* indicate an excess of rare haplotypes compared to those expected for a stable population, indicating a possible population expansion (Fu 1997). Large significant positive values of *Fs* indicate a greater deficit of rare haplotypes compared to those expected for a stable population indicating that the population has probably experienced a bottleneck (Tajima 1989; Fu 1997). Statistical significance was obtained after 1,000 simulations of *Fs* values under the null hypothesis of population equilibrium using a coalescent approach (Schneider *et al.* 2000). As other evolutionary processes such as genetic hitch-hiking and background selection can leave similar imprints to those of population expansions, we also calculated the statistics F* and D* (Fu & Li 1993) using the program *ProSeq* version 2.9 (Filatov 2002), with statistical significance obtained after 1,000 simulations. If *Fs* is negative and significant but F* and D* are not significant, then a population expansion is indicated. In contrast, if *Fs* is not significant and F* and D* are significant then background selection is more likely to have cause the observed pattern (Fu 1997).

5.5 Results

5.5.1 MtDNA control region sequence characteristics

Of the initial 695 individuals, 625 were successfully sequenced for ≥ 821bp of the mtDNA control region and therefore were included in the analyses. The alignment of the 625 sequences resulted in 74 haplotypes, with 57 polymorphic sites of which 28 were parsimony informative. Average number of substitutions between haplotypes was low ($k = 5.456$).

5.5.2 Phylogenetic analyses

Phylogenetic relationships of the haplotypes and the sequences downloaded from Genbank are illustrated in Figure 5.2. Only the phylogenetic un-rooted tree generated with Bayesian inference is shown as it presented a more resolved topology. Posterior probability support values from the Bayesian inference (BPP) are shown at the top of the branches and maximum parsimony bootstrap values (MPB) are shown below branches. All the species within the genus *Cervus*, except *C. eldi thamin* which is sometimes placed within other genera *Rucervus* or *Panolia* (Groves 2006), formed a well supported group (100 BPP, 80.5 MPB), sister to the more distant taxa of *Dama mesopotamica*, *C. eldi thamin, Axis porcinus* and *Elaphurus davidianus*. Phylogenetic relationships between *Cervus* and the latter four taxa were only resolved by Bayesian inference but with very low posterior probability support. The monophyletic clade containing all the *Cervus* species (except *C. eldi thamin*) was divided into two well supported clades, one containing sika deer (*C. nippon*), wapiti (*C. canadensis* or *C. elaphus canadensis*) and Thorold's white-lipped deer (*C. albirostris*) (100 BPP, 85.3MPB) and another containing red deer (99 BPP, 57.3 MPB). Within the wapiti/sika/white-lipped deer clade, wapiti formed a separate group (100 BPP, 90 MPB) and sika deer clustered with Thorold's white-lipped deer (84 BPP, 68.2 MPB). The wapiti clade was resolved in the Bayesian inference but maximum parsimony analyses only supported the relationships between *C. c. wallichi* and *C. c. alashanicus* (100 BPP, 90 MPB), and relationships between *C. c. songaricus* and *C. c. xanthopygus* (91 BPP, 100MPB). However, Bayesian inference

provided a high posterior probability (92 BPP) for the group containing all North American wapiti. The red deer clade consisted of two sister lineages, one including Central Asian red deer (*C. e. yarkandensis* and *C. e. bactrianus*) (99.9 BPP, 100 MPB) and another lineage containing all European and North African red deer (92.2 BPP, 100MPB). Within the European-North African polytomy, North African red deer (*C. e. barbarus*), Corsican red deer (*C. e. corsicanus*) and the haplotype Rum A found in the island of Rum (Scotland) formed a very well supported group (100 BPP, 99.1 MPB) and the Rum A haplotype clustered with Corsican red deer (100 BPP, 98.5 MPP). The inclusion of all Western European red deer in a separate clade was only supported in the Bayesian inference (97 BPP). Scandinavian red deer (*C. e. atlanticus*), Spanish red deer (*C. e. hispanicus*), all other haplotypes found in studies of British red deer and all haplotypes found in this study fell within the Western European red deer polytomy.

Despite the lack of high parsimony bootstrapping support between haplotypes within the Western European red deer polytomy, visual inspection of the tree suggests some levels of geographical structure as particular haplotypes were found in particular areas of the Scottish Highlands (Fig. 5.2). In the tree some groups of haplotypes only found in a particular population (or adjacent populations) have been highlighted in different colours to illustrate local haplotype structure. Thick black lines highlight possible translocation events in the Scottish Highlands. For example H18 (found in 1 female in CON and 1 male in AG) clustered with haplotypes found on the island of Arran (SW of Scotland) and Euston (Suffolk, England) suggesting that this haplotype either originated from a translocation or introduction event from the same stock, or the haplotypes found in Arran and Euston originated from mainland Scotland, probably areas close to CON-AG (Fig. 5.1 & 5.2). Another translocation or introduction event was also suggested for H16 (found in 2 males in CON and 1 male and 1 female in AG) that clustered with haplotypes found in Euston, Arran and Rum.

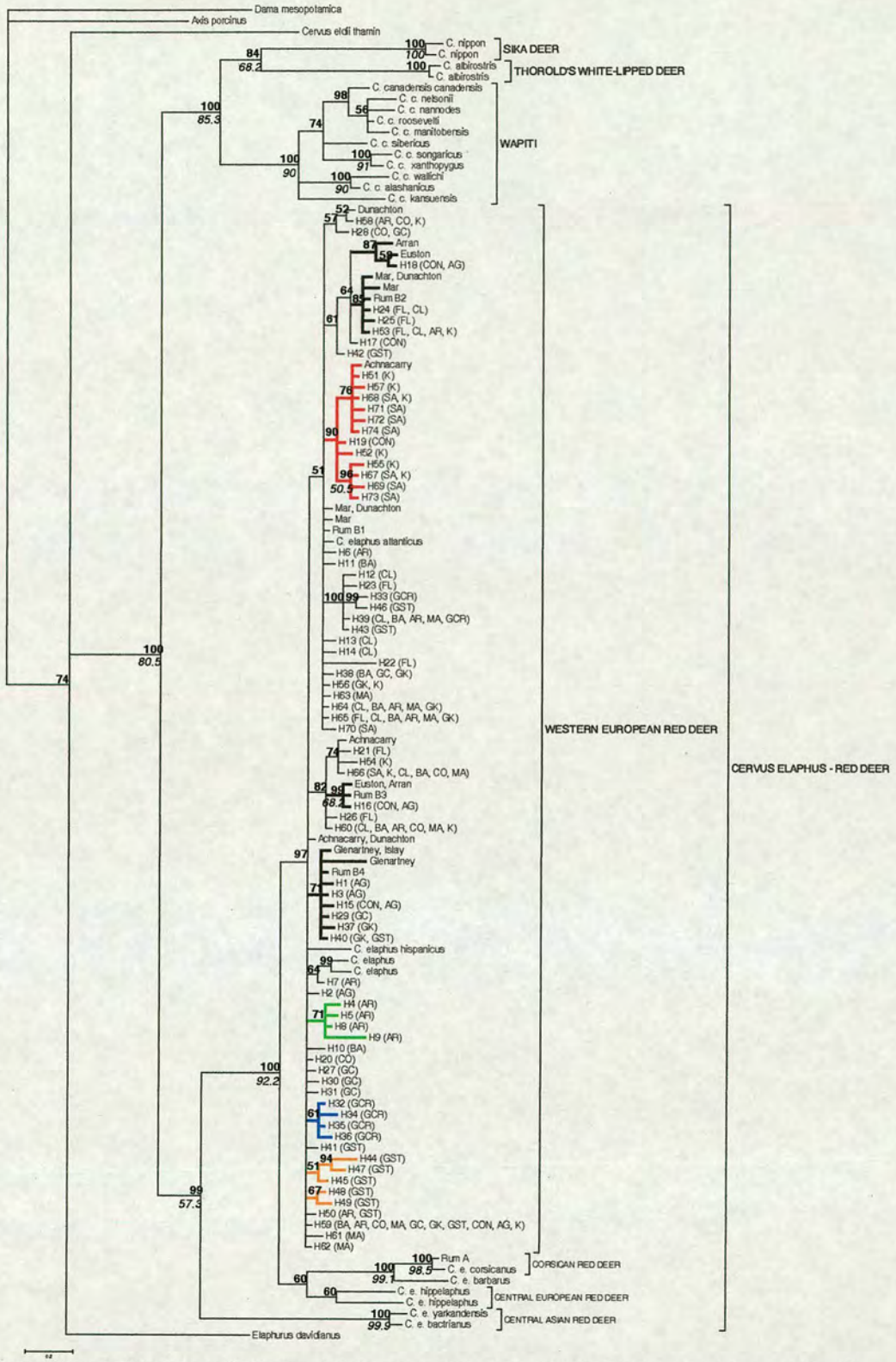


Fig. 5.2 Unrooted phylogram illustrating phylogenetic relationships of Scottish Highland red deer and the 47 sequences of Cervinae download from Genbank. The phylogram was constructed using Bayesian inference with a GTR + G model (Rodríguez *et al.* 1990) for mtDNA control region sequences. Bayesian posterior probability values are shown in bold and above the branches. Maximum parsimony bootstrap support values (Felsenstein 1985) are show in italics and below branches. Branch length units are expected substitutions per site. Haplotypes are labelled as species or locations of haplotypes found in previous studies or as labelled in this study (see Table 5.2 & 5.3). For those haplotypes found in this study, the sampling sites where they were found are show in brackets (see Fig. 5.1 legend for sampling area codes). Coloured branches represent haplotypes only found in a particular population (or adjacent populations) to illustrate local population structure. Branches with thick black lines highlight possible red deer translocation events.

5.5.3 Population analyses

5.5.3.1 Genetic diversity

Scottish Highland red deer presented high levels of mtDNA CR diversity, with 74 haplotypes identified in a total of 625 individuals. Genetic diversity indices are summarised in Table 5.2. AR presented the highest haplotype diversity (0.8364 ± 0.0262) and GCR the lowest (0.2529 ± 0.1037). The highest nucleotide diversity was found in K (0.0057 ± 0.0031) and the lowest in GCR (0.0006 ± 0.0006). The distribution of haplotypes is illustrated in Table 5.3. Some haplotypes were common and shared between populations, with H59 being the most common haplotype, found in 25.92% of the individuals. H59 was found at a relatively high frequency in most of the populations except in FL, CL, GCR, SA where it was absent and in K where only one individual presented H59. Other haplotypes were common in particular adjacent populations of the study area (e.g. H15 in CON and AG, H66, 67, 68 in SA and K) or confined to individual populations (e.g. H32 in GCR). Haplotypes unique to one individual represented 7.2% of the sample. When analysing the data for each sex separately, similar patterns of haplotype distribution were found (data not shown). Haplotype H59 was the most common and widely distributed haplotype in both sexes (28.12% of the males and 24.37% of the females). Haplotypes unique to one individual were present in 7.17% and 10.72% of the male and female datasets, respectively.

Table 5.2. Genetic diversity analyses from the 74 haplotypes found in 14 populations of red deer in the Scottish Highlands

Population	no. individuals	no. haplotypes	Haplotype diversity $h \pm SD$	Nucleotide diversity $\pi \pm SD$
FL	41	8	0.7059 ± 0.0663	0.0036 ± 0.0021
CL	60	10	0.8328 ± 0.0256	0.0055 ± 0.003
BA	42	9	0.7875 ± 0.0422	0.0033 ± 0.002
AR	55	13	0.8364 ± 0.0262	0.0035 ± 0.002
CO	27	7	0.5983 ± 0.1048	0.0027 ± 0.0017
MA	51	8	0.5090 ± 0.0779	0.0018 ± 0.0012
GC	49	8	0.6420 ± 0.0602	0.0025 ± 0.0016
GCR	30	5	0.2529 ± 0.1037	0.0006 ± 0.0006
GK	49	7	0.6998 ± 0.0418	0.0031 ± 0.0019
GST	34	12	0.7576 ± 0.0570	0.0034 ± 0.0021
CON	48	6	0.5621 ± 0.0682	0.0037 ± 0.0022
AG	25	7	0.6967 ± 0.0647	0.0037 ± 0.0022
SA	58	9	0.7417 ± 0.0255	0.0056 ± 0.0031
K	55	13	0.8061 ± 0.0335	0.0057 ± 0.0031

Table 5.3. Distribution of Scottish Highland red deer mtDNA CR haplotypes. Each column shows the number of individuals from a population which had a particular haplotype. Numbers in parenthesis are total sample sizes for each population.

	Populations (n)													
Haplotypes	FL (41)	CL (60)	BA (42)	AR (55)	CO (27)	MA (51)	GC (49)	GCR (30)	GK (49)	GST (34)	CON (48)	AG (25)	SA (58)	K (55)
H1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
H2	0	0	0	0	0	0	0	0	0	0	0	1	0	0
H3	0	0	0	0	0	0	0	0	0	0	0	1	0	0
H4	0	0	0	1	0	0	0	0	0	0	0	0	0	0
H5	0	0	0	2	0	0	0	0	0	0	0	0	0	0
H6	0	0	0	1	0	0	0	0	0	0	0	0	0	0
H7	0	0	0	1	0	0	0	0	0	0	0	0	0	0
H8	0	0	0	1	0	0	0	0	0	0	0	0	0	0
H9	0	0	0	1	0	0	0	0	0	0	0	0	0	0
H10	0	0	1	0	0	0	0	0	0	0	0	0	0	0
H11	0	0	2	0	0	0	0	0	0	0	0	0	0	0
H12	0	1	0	0	0	0	0	0	0	0	0	0	0	0
H13	0	2	0	0	0	0	0	0	0	0	0	0	0	0
H14	0	1	0	0	0	0	0	0	0	0	0	0	0	0
H15	0	0	0	0	0	0	0	0	0	0	30	9	0	0
H16	0	0	0	0	0	0	0	0	0	0	3	1	0	0
H17	0	0	0	0	0	0	0	0	0	0	1	0	0	0
H18	0	0	0	0	0	0	0	0	0	0	1	1	0	0
H19	0	0	0	0	0	0	0	0	0	0	2	0	0	0
H20	0	0	0	0	1	0	0	0	0	0	0	0	0	0
H21	3	0	0	0	0	0	0	0	0	0	0	0	0	0
H22	1	0	0	0	0	0	0	0	0	0	0	0	0	0
H23	2	0	0	0	0	0	0	0	0	0	0	0	0	0

H24	4	6	0	0	0	0	0	0	0	0	0	0	0	0
H25	2	0	0	0	0	0	0	0	0	0	0	0	0	0
H26	1	0	0	0	0	0	0	0	0	0	0	0	0	0
H27	0	0	0	0	0	0	1	0	0	0	0	0	0	0
H28	0	0	0	0	2	0	6	0	0	0	0	0	0	0
H29	0	0	0	0	0	0	1	0	0	0	0	0	0	0
H30	0	0	0	0	0	0	1	0	0	0	0	0	0	0
H31	0	0	0	0	0	0	1	0	0	0	0	0	0	0
H32	0	0	0	0	0	0	0	26	0	0	0	0	0	0
H33	0	0	0	0	0	0	0	1	0	0	0	0	0	0
H34	0	0	0	0	0	0	0	1	0	0	0	0	0	0
H35	0	0	0	0	0	0	0	1	0	0	0	0	0	0
H36	0	0	0	0	0	0	0	1	0	0	0	0	0	0
H37	0	0	0	0	0	0	0	0	2	0	0	0	0	0
H38	0	0	2	0	0	0	1	0	1	0	0	0	0	0
H39	0	12	3	0	1	0	11	0	3	0	0	0	0	0
H40	0	0	0	0	0	0	0	0	20	10	0	0	0	0
H41	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H42	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H43	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H44	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H45	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H46	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H47	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H48	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H49	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H50	0	0	0	2	0	0	0	0	0	1	0	0	0	0
H51	0	0	0	0	0	0	0	0	0	0	0	0	0	1
H52	0	0	0	0	0	0	0	0	0	0	0	0	0	1
H53	21	18	0	1	0	0	0	0	0	0	0	0	0	1
H54	0	0	0	0	0	0	0	0	0	0	0	0	0	1
H55	0	0	0	0	0	0	0	0	0	0	0	0	0	1
H56	0	0	0	0	0	0	0	0	1	0	0	0	0	1
H57	0	0	0	0	0	0	0	0	0	0	0	0	0	2
H58	0	0	0	2	1	0	0	0	0	0	0	0	0	1
H59	0	0	15	13	17	35	27	0	18	14	11	11	0	1
H60	0	1	12	13	2	8	0	0	0	0	0	0	0	4
H61	0	0	0	0	0	1	0	0	0	0	0	0	0	0
H62	0	0	0	0	0	1	0	0	0	0	0	0	0	0
H63	0	0	0	0	0	1	0	0	0	0	0	0	0	0
H64	0	4	4	13	0	1	0	0	0	0	0	0	0	0
H65	7	10	2	4	0	1	0	0	4	0	0	0	0	0
H66	0	5	1	0	3	3	0	0	0	0	0	0	17	11
H67	0	0	0	0	0	0	0	0	0	0	0	0	16	11
H68	0	0	0	0	0	0	0	0	0	0	0	0	19	19
H69	0	0	0	0	0	0	0	0	0	0	0	0	1	0
H70	0	0	0	0	0	0	0	0	0	0	0	0	1	0
H71	0	0	0	0	0	0	0	0	0	0	0	0	1	0
H72	0	0	0	0	0	0	0	0	0	0	0	0	1	0
H73	0	0	0	0	0	0	0	0	0	0	0	0	1	0
H74	0	0	0	0	0	0	0	0	0	0	0	0	1	0

Relationships between the haplotypes are illustrated in the minimum spanning network (Fig. 5.3). Haplotype H59 appeared at one extreme of the network giving rise to the rest of haplotypes. Haplotype structuring was evident due to the fact that some star-like phylogenies were found in different parts of the network. As already suggested in the phylogenetic tree, some of the star-like groups were found only in a particular population or in closely located populations which might represent evolution *in situ*, with related haplotypes radiating from a more common one (e.g. cluster around H32 found only in GCR, the clusters around H67 and H68 found only SA-K).

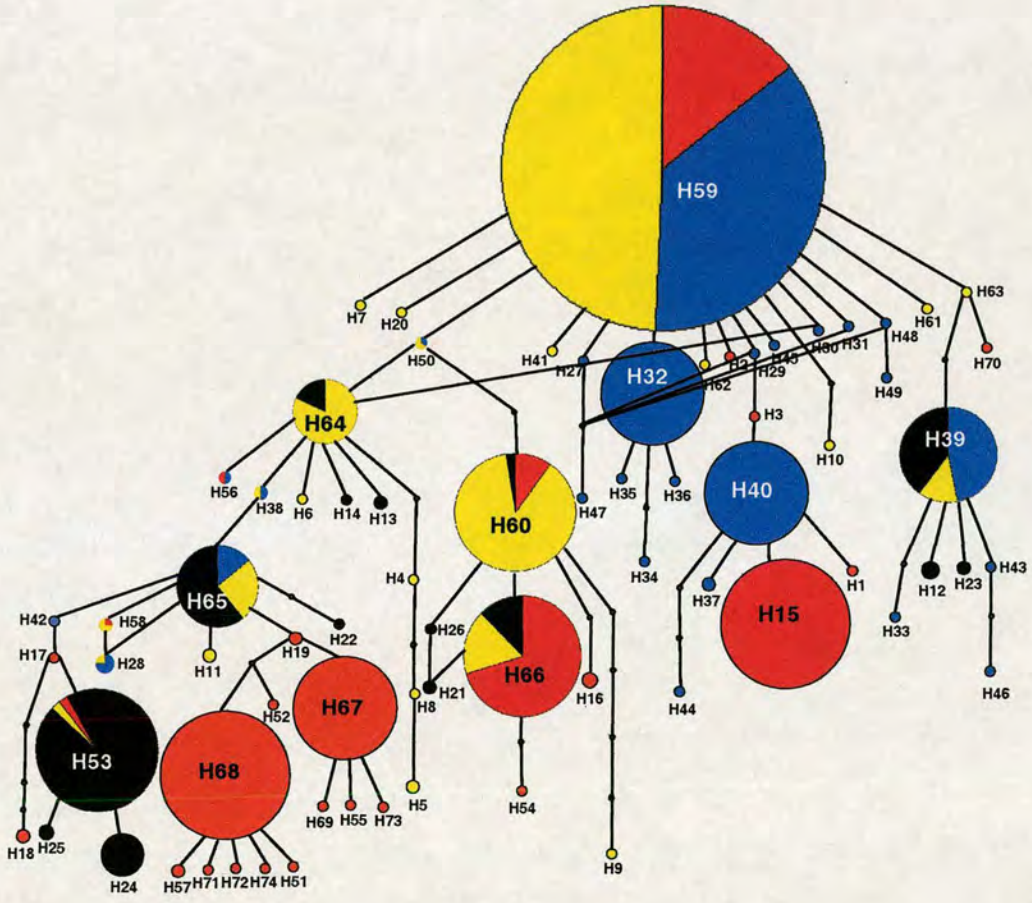


Fig. 5.3 Minimum spanning network based on the number of substitution differences among Scottish Highland red deer haplotypes. Circle size is scaled to the number of individuals presenting that particular haplotype. Colours indicate the proportion of individuals sampled in each of the four regions identified by microsatellite analyses (see Chapter 3; Pérez-Espona *et al.* submitted) in the study area that presented that particular haplotype. Branches are not scaled to the number of nucleotide substitutions. Branches without nodes represent one nucleotide difference, and nodes represent additional nucleotide differences.

5.5.3.2 Population structure and sex-biased dispersal

Significant population genetic structure was found in the study area (global Φ_{ST} = 0.3483, P < 0.001). Pair-wise population differentiation tests (after applying the Bonferroni correction) showed that most of the populations were significantly differentiated from each other except for the following adjacent or closely located populations: FL-CL, BA-AR, AR- CO, BA-CO, BA-MA, CO-MA, CO-GC, GK-GST, GK-AG, GST-AG, CON-AG and SA-K (Table 5. 4). Estimates of Φ_{ST} across the study area were similar for both sexes ($\Phi_{STmales}$ = 0.3817, P < 0.01; $\Phi_{STfemales}$ = 0.3698, P < 0.01) indicating that male-biased dispersal is not likely to affect mtDNA CR population structure. Comparisons of pair-wise population differentiation measures for each sex separately further confirmed that male-biased dispersal is not strong in the study area. Evidence of strong male-biased dispersal was only suggested between the adjacent estates of FL and CL as these populations were not significantly differentiated from each other when using the male data set but significantly differed using the female dataset (Table 5.5). Pair-wise population comparisons in fact suggested female-biased dispersal between several non-adjacent estates (BA-CL, AR-CL, BA-GC, AR-GC, AR-GST, BA-SA, BA-K) as they could not be differentiated from each other when using the female data set but were significantly differentiated when using the male data set (Table 5.5). To assess the impact of the lack of female samples from CO into the pair-wise measures population differentiation test for the male data set were repeated excluding samples from CO but patterns of differentiation did not differ (data not shown). Hierarchical AMOVAs showed that 18.25% of the total genetic variation was attributable to variation among the four regions previously identified by microsatellite analyses (Chapter 3; Pérez-Espona *et al.* submitted): East (FL, CL), Central 1 (AR,BA,CO,MA), Central 2 (GC, GCR, GK, GST) and West (CON, AG, SA, K) (Table 5.6). Hierarchical AMOVAs conducted for each sex separately showed that there were higher levels of population structuring among regions for males than for females ($\Phi_{CTmales}$ = 0.238, P < 0.0001; $\Phi_{CTfemales}$ = 0.126, P = 0.033). In contrast, female population structure within regions was higher than males ($\Phi_{SCmales}$ = 0.188, P < 0.001; $\Phi_{SCfemales}$ = 0.279, P < 0.001).

Table 5.4. Pair-wise population differentiation estimates. The below diagonal matrix gives values of mtDNA Φ_{ST}, the above diagonal matrix gives values of microsatellite F_{ST}. Bold indicates population pairs that were not significantly differentiated after applying Bonferroni correction.

Population	FL	CL	BA	AR	CO	MA	GC	GCR	GK	GST	CON	AG	SA	K
FL		**0.0026**	0.0188	0.0156	0.0188	0.0193	0.0272	0.0322	0.0273	0.032	0.0273	0.0319	0.0275	0.0309
CL	**0.1107**		0.0163	0.0122	0.0109	0.0146	0.021	0.0297	0.0204	0.0259	0.0286	0.0338	0.0308	0.0334
BA	0.5276	0.2627		**0**	**0.0031**	0.0081	0.0106	0.0126	0.0083	0.0093	0.0242	0.0298	0.0251	0.0262
AR	0.4896	0.2371	**0.0081**		**0.0005**	**0.0015**	0.0084	0.0142	0.0052	0.0051	0.0233	0.0274	0.0219	0.025
CO	0.5737	0.3062	**0.0296**	**0.0696**		**0.0052**	0.0095	0.0115	0.004	**0.0061**	0.0231	0.0317	0.0236	0.0215
MA	0.6629	0.4081	**0.0750**	0.1266	**-0.0051**		0.0059	0.0128	0.0062	0.0053	0.0196	0.0253	0.0226	0.0227
GC	0.5947	0.3157	0.1340	0.1725	**0.0729**	0.1470		0.0054	**0.0023**	0.0038	0.0328	0.0421	0.0309	0.0335
GCR	0.7476	0.5234	0.4606	0.4579	0.4666	0.5057	0.4451		0.0064	0.0042	0.0327	0.0434	0.0295	0.0339
GK	0.5796	0.3449	0.1769	0.1751	0.1754	0.2331	0.2190	0.4916		0.0027	0.0319	0.0411	0.0317	0.0348
GST	0.5866	0.3499	0.1417	0.1595	**0.0953**	0.1356	0.1399	0.4299	**0.0202**		0.0301	0.0326	0.0267	0.0306
CON	0.5997	0.4198	0.3192	0.3091	0.3345	0.3950	0.4006	0.5631	0.1459	0.1933		**0**	0.0066	0.0064
AG	0.5931	0.3723	0.2051	0.2111	0.1828	0.2400	0.2567	0.5096	**0.0524**	**0.0498**	**0.0393**		0.009	0.0095
SA	0.3762	0.2391	0.2879	0.2842	0.3245	0.4090	0.3822	0.5322	0.38256	0.3815	0.4455	0.3982		**0.0022**
K	0.3605	0.2146	0.2572	0.2528	0.2938	0.3793	0.3520	0.5128	0.3574	0.3545	0.4271	0.3739	**-0.0112**	

Table 5.5. Comparisons of pair-wise population differentiation estimates between the sexes using mtDNA Φ_{ST}. Above diagonal, estimates obtained for females. Below diagonal, estimates obtained for males. In bold, populations that were not significantly differentiated after applying Bonferroni correction.

Population	FL	CL	BA	AR	CO	MA	GC	GCR	GK	GST	CON	AG	SA	K
FL		0.1930	0.4924	0.4700	-	0.7034	0.5722	0.7000	0.6226	0.6242	0.6330	0.62671	0.38806	0.38308
CL	**0.0223**		**0.1316**	**0.1329**	-	0.3713	0.2061	0.4066	0.3348	0.3136	0.4172	0.34718	0.19043	0.18072
BA	0.5367	0.3728		**0.0175**	-	**0.1492**	**0.0879**	0.3612	0.2337	**0.1889**	0.3602	0.20712	**0.20364**	**0.1828**
AR	0.5078	0.3620	**-0.0084**		-	0.1850	**0.1042**	0.3845	0.2255	**0.1938**	0.3584	0.22263	0.24601	0.2282
CO	0.5707	0.4094	**0.0192**	**0.0815**		-	-	-	-	-	-	-	-	-
MA	0.6111	0.4473	**-0.0006**	**0.0527**	**-0.0055**		0.1563	0.5031	0.3633	**0.3112**	0.4814	**0.2681**	0.4221	0.4023
GC	0.6111	0.4259	**0.1504**	0.2223	**0.0638**	**0.1528**		0.3546	0.3000	**0.2478**	0.4316	0.2581	0.3544	0.3293
GCR	0.7754	0.6162	0.5030	0.4967	0.4495	0.5235	0.5027		0.5302	0.5428	0.5576	0.4525	0.4696	0.4487
GK	0.5404	0.3911	0.1412	0.1653	0.1413	0.1778	0.1770	0.4742		**-0.0662**	**0.1494**	**0.0600**	0.3840	0.3705
GST	0.5548	0.4102	0.1143	0.1594	**0.0720**	**0.1061**	**0.0928**	0.3994	**0.0193**		**0.1535**	**0.0040**	0.3710	**0.3524**
CON	0.5438	0.4215	0.2352	0.2289	0.2607	0.2706	0.3389	0.5510	**0.0957**	**0.1421**		**0.0732**	0.4662	0.4588
AG	0.5213	0.3706	**0.1782**	**0.1876**	**0.1732**	**0.2067**	**0.2409**	0.5584	**0.0112**	**0.0322**	**-0.0216**		0.3960	0.3785
SA	0.3733	0.3052	0.3718	0.3723	0.3988	0.4236	0.4286	0.5942	0.4097	0.4128	0.4198	0.3967		**-0.02715**
K	0.3273	0.2506	0.3090	0.3103	0.3390	0.3639	0.3734	0.5538	0.3544	0.3599	0.3726	0.3445	**-0.0200**	0

Table 5.6. Results from hierarchical AMOVAs. 'Regions' are the four areas identified by previous microsatellite analyses. All Φ statistics were significant (P < 0.05).

	All	Males	Females
Percentage of variance components			
Among regions	18.25	23.28	12.57
Among populations within regions	18.96	14.34	24.42
Within populations	62.79	61.83	63.01
Φ statistics			
Φ_{CT}	0.183	0.238	0.126
Φ_{SC}	0.24	0.188	0.279
Φ_{ST}	0.372	0.382	0.37

5.5.3.3 Population expansion analyses

Mismatch distribution analyses showed a unimodal distribution of pair-wise nucleotide differences indicated a significant population expansion when considering the study area as a single population (Fig. 5.4). The raggedness index and the goodness-of-fit tests suggested a population expansion (SSD = 0.005, P = 0.482; *rg* = 0.015, P = 0.583). Fu's *Fs* neutrality tests also suggested a possible population expansion (*Fs* = -24.94; P < 0.0001), and background selection as the cause for the observed genetic pattern was discarded by Fu & Li's tests (D* = -3.09; P > 0.05; F* = -2.62, P > 0.05).

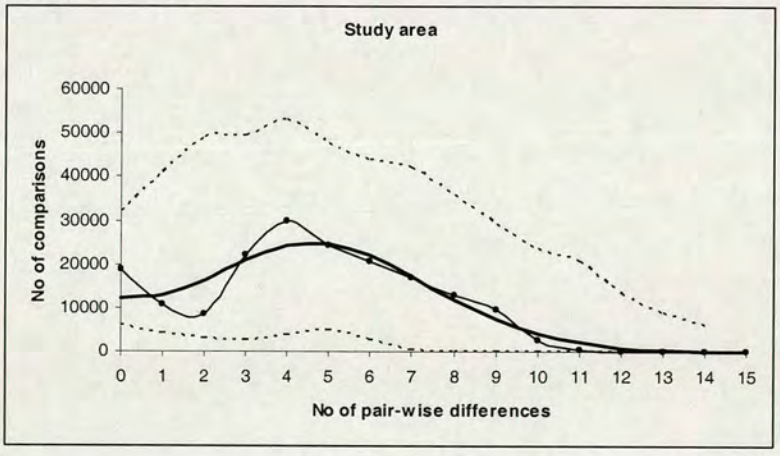


Fig. 5.4 Mismatch distribution of pair-wise nucleotide differences between 625 mtDNA control region sequences of Scottish Highland red deer. Observed distribution of pair-wise nucleotide differences (thin line). Distribution of pair-wise nucleotide differences expected under a model of population expansion (thick line). Upper and lower bound of the 95% confidence intervals for the observed values (dashed lines).

The timing of the population expansion in units of mutation (τ) was 5.547. The timing of the population expansion in years (t) can be estimated from $\tau = 2ut$, where $u = \mu k$, μ= mutation rate and k = length of the sequence. Considering an approximate mutation rate for the control region of 0.036 substitutions/bp/Mya (Polziehn & Strobeck 2002), the estimated date for the population expansion is c. 93,000 years BP (95% CI 44,000 – 145,774 years BP).

5.6 Discussion

5.6.1 Genetic integrity of Scottish Highland red deer and phylogenetic relationships with other Cervus species

Despite the documented introductions and interbreeding of Scottish red deer with foreign species such as sika deer and wapiti (Whitehead 1960, 1964; Clutton-Brock & Albon 1989), none of the individuals included in this study showed introgression of mtDNA haplotypes from other subspecies or species of deer. Due to the large number of sequences included in this study, the lack of individuals with introgressed mtDNA haplotype suggests that unlike for the red deer population in the island of Rum where a haplotype of Corsican deer was found (Nussey *et* al. in press) or Argyll where hybridisation between sika and red deer has been detected (Abernethy 1994, Goodman *et al.* 1999), the levels of successful introduction and hybridisation between red deer and foreign species or subspecies in the area of the Scottish Highlands sampled has been limited. Hybridisation of red deer with the much more closely related Spanish and Norwegian red deer can also be discarded as although phylogenetic relationships between these subspecies can not be resolved with mtDNA control region (see also Randi *et al.* 2001) or cytochrome b (Ludt *et al.* 2004; Pitra *et al.* 2004), none of the haplotypes found in the Scottish Highlands clustered tightly with either the Norwegian or the Spanish haplotypes. However, due to the maternal inheritance of mitochondria, hybridisation events involving males of foreign subspecies or species of deer with red deer females will remain undetected. Introductions of foreign stocks in Scotland were aimed at improving of hunting

trophy quality and consisted predominantly of males (Whitehead 1960, 1964). Therefore, a male-specific marker in the Y-chromosome with variation between species and subspecies would be needed to thoroughly assess the extent of successful hybridisation between foreign sub/species of deer and Scottish Highland red deer.

Although we did not find any individual with introgressed mtDNA from other subspecies or species of deer, inspection of the phylogenetic relationships between haplotypes found in our study area and those found in other British locations (Hwme *et al.* in press; Nussey *et al.* in press) suggested a few possible translocation events. Two haplotypes found in very few individuals in the estates of CON and AG clustered with haplotypes found in the Scottish islands (Arran, Islay and Rum) and Euston (England), locations for which part or the entire population of red deer is known to have originated from introductions events of individuals originated in English deer parks as well as mainland Scotland (Hmwe *et al.* in press, Nussey *et al.* in press). Documented introductions of deer park blood into the estates included in our study area are only known for AG where one stag from Warnham Court Park was introduced in 1920 (list of imports and exports from Warnham Court Park, unpublished data) and MA, where several wapiti individuals were introduced in the 19th Century. However, translocation events of individuals from deer parks could have taken place on estates nearby to those included in this study and subsequently spread through female dispersal. Some haplotypes found on Rum did cluster with some of our samples, suggesting that they might have originated from mainland Scotland populations. Haplotypes found in mainland Scotland by Hmwe *et al.* (in press) clustered with haplotypes sampled from geographically close areas included in our study. This suggests that human activities have not completely blurred red deer population boundaries in the Scottish mainland

5.6.2 Genetic diversity and population structure of red deer in the Scottish Highlands

Scottish highland red deer showed high levels of genetic diversity at the mtDNA control region. There was moderate to high average haplotype diversity (0.6735) and low average nucleotide diversity (0.0035) as expected due to the low divergence

between haplotypes at the scale at which this study was conducted. Relatively high haplotype diversity and low nucleotide diversity would indicate that several mtDNA lineages have evolved yielding a relatively high number of haplotypes differing by a small number of bases (average 5.65 substitutions, maximum 8 substitutions between H59 and H9). Direct comparisons of the genetic diversity found for Scottish Highland red deer with other studies of deer or ungulates, or in fact to other mammals, are difficult as no other study of mammals has assessed so many individuals for a mtDNA regional study as the one presented here. Despite these differences in sampling regime, genetic diversity found in the Scottish Highland red deer populations was within the range of that found in other studies of ungulates in which a more sparse sample collection was undertaken (e.g. Ramey 1995; Nagata *et al.* 1998; Boyce *et al.* 1999; Vernesi *et al.* 2002; Hartl *et al.* 2003; Zachos *et al.* 2003; Feulner *et al.* 2004; Randi *et al.* 2004; Sarno *et al.* 2004). Genetic diversity in Scottish highland red deer was also much higher than that found on the Scottish islands (Hmwe *et al.* in press; Nussey *et al.* in press), with the exception of GCR which presented low genetic diversity values (h = 0.2529) similar to those found in Arran (Hmwe *et al.* in press). The low genetic diversity in GCR can be explained by the small population of red deer that is found in the estate (Whitehead 1960, 1964) and also by the fact that is currently surrounded by fenced forests which might prevent red deer gene flow into GCR (Chapter 3; Pérez-Espona *et al.* submitted).

The wide distribution and high frequency of haplotype H59 in the minimum spanning network, giving rise to the rest of the haplotypes found in the study area suggests that this haplotype might represent an ancestral population that subsequently expanded (Donnelly & Tavaré 1986; Takahata 1998; Crandall & Templeton 1993). The wide distribution of H59 in the study area also reflects the high levels of gene flow among populations located in the central regions of the study area (Lavery *et al.* 1996; Escorza-Treviño & Dizon 2000). Mismatch distribution analysis and neutrality tests confirmed a population expansion and the fact that the peak in the mismatch distribution was at low sequence divergence indicates that the population of red deer in the Scottish Highlands is relatively new and that due to a population expansion there was a gradual accumulation of

nucleotide differences between those haplotypes derived from the ancestral one (Harpending *et al.* 1993). Population expansions detected at the relatively small scale of our study area can be linked to colonisation processes (Endler 1977). Rare long-distance dispersal events undertaken by individuals from the ancestral population might have created new populations producing a cluster of related haplotypes specific to a geographical area that can persist for hundreds of generations or longer due to the presence of gene flow barriers (Ibrahim *et al.* 1996). Colonisation of new areas in the Scottish mainland during interglacial periods might have resulted in the evolution of new derived haplotypes from the putative ancestral H59 which due to limited historical and contemporary gene flow between regions plus the effects of genetic drift might have resulted in haplotypes being restricted to particular areas (e.g. haplotypes H66, 67, 68 in SA and K; H32 in GCR). Although an accurate estimate of the mutation rates for the mtDNA control region for red deer is not available and therefore estimates of the time of the population expansion which rely on the mutation rate might not be accurate, using the approximate mutation rate of 0.036 substitutions/bp/Mya (Polziehn & Strobeck 2002) the time of the population expansion was estimated to be c. 93,000 years BP. This date suggests that the population expansion detected in this study is not reflecting population size fluctuations during the past centuries but rather reflects a more historical growth that could have occurred during the last interglacial (Ipswichian), c. 125,000-110,000 years BP. Red deer were found in Britain during most of this period (Lister 1984). However, red deer is known to have been absent from Scotland during the last glaciations and colonised Scotland after the ice sheets retreated (c. 11,0000 years BP) (Lister 1984; Clutton-Brock 1989) and therefore, we would have expected that the population expansion would have reflected this colonisation after the glaciations. Further studies aiming at an accurate estimation for the mutation rate of the mitochondrial control region of red deer will help to obtain a more reliable date for the suggested population expansion.

Despite the relatively small scale of the study area, population structure analyses revealed significant mtDNA differentiation across the study area ($\Phi_{ST} = 0.3483$), with significant differentiation among the four regions previously identified by

microsatellite analyses (Chapter 3; Pérez-Espona *et al.* submitted) (Table 5.6). Pair-wise population differentiation analyses yielded similar results to those obtained with microsatellite markers with most adjacent or nearby populations not being significantly differentiated from each other (Table 5.4). However, there were two pairs of estates sampled on either side of the Great Glen (GK-AG, GST-AG) that were not significantly differentiated from each other when using mtDNA control regions but presented significant and relatively large F_{ST} estimates when using microsatellite markers (Table 5.4). Although translocation events could explain the lack of differentiation between these two pairs of populations for mtDNA CR, the only haplotype shared between AG-GK and AG-GST was H59, the most common and putative ancestral haplotype found in the study area. In this case, it suggested that the lack of differentiation is due to gene flow rather than translocations. Discrepancy between mtDNA and microsatellites would be more likely as a result of decreased contemporary gene flow between these populations due to the presence of current man-made landscape features acting as red deer gene flow barriers such as roads and fenced forests (Chapter 3; Pérez-Espona *et al.* submitted). The effect of contemporary gene flow barriers along the Great Glen on Scottish Highland red deer population structure was also suggested by the fact that the largest pair-wise Φ_{ST} estimates were found between between distant populations on the same side of the Great Glen (FL-GCR) in contrast with results obtained with microsatellite markers in which the largest F_{ST} was found between closely located populations located at either side of the valley (AG-GCR).

5.6.3 Sex-biased dispersal

Population structure estimates obtained with mtDNA (Φ_{ST} = 0.3483) were much larger than the four-fold expected between mtDNA and microsatellite markers (F_{ST} = 0.019). Therefore, comparison of the two estimates would suggest a strong male-biased dispersal in the study area (e.g. Fitzsimmons *et al.* 1997; Lyrholm *et al.* 1999; Escorza-Treviño & Dizon 2000; Kerth *et al.* 2002). However, the levels of population differentiation in terms of F_{ST} values obtained from microsatellite data might be underestimated due to the high polymorphism of the microsatellite markers

(Hedrick 1999; Balloux & Lugon-Moulin 2002). In fact, when using the standardised measure of population differentiation (G_{ST}') proposed by Hedrick (2005), much larger levels of microsatellite differentiation were obtained ($G_{ST}' = 0.14$) (Chapter 3; Pérez-Espona *et al.* submitted). If this estimate is compared to the one obtained with mtDNA CR sequences, the suggestion of strong male-biased dispersal is no longer valid and the difference between markers is in fact slightly smaller than the expected four-fold difference. However, a four-fold difference between the two kinds of molecular markers are expected under the assumptions of an ideal population with random mating, conditions which are often not encountered in natural populations (Chesser & Baker 1996). Breeding systems, dispersal behaviour of both sexes and levels of population differentiation might lead to different expectations from the four-fold difference between mtDNA and microsatellites (Chesser & Baker 1996). For polygynous mammals such as red deer, where only few males contribute to the next generation and females are generally more philopatric than males, effective population sizes of mtDNA are expected to be much smaller than the four-fold difference (Chesser & Baker 1996) and if the effective number of males is less than one seventh of the number of effective females, genetic drift in nuclear genes might be higher than for mitochondrial genes (Ballard & Whitlock 2004). The mitochondrial differentiation estimate obtained in our study is about 2.4 times larger than the one obtained with microsatellites (considering the G_{ST}' value), suggesting that there is male-biased dispersal across the study area but that the bias is not extreme. It is important to note that due to the generally non-recombining nature of mtDNA (although see Ladoukakis & Zouros 2001), selective sweeps (the fixation of one haplotype due to high fitness) or background selection (reductions in effective population sizes due to the elimination of low fitness haplotypes) will affect the levels of population structure observed (Ballard & Whitlock 2004). Background selection was discarded by the Fu & Li (1993) analyses conducted in this study, the effects of selective sweeps (genetic hitchhiking) leave a similar genetic signature as population expansion events. However, as patterns of population structure found with mtDNA CR sequences were similar to those found with microsatellite markers, it is unlikely that selection influenced both types of genetic markers in a similar way or that a selective sweep occurred at the mtDNA control region without altering the

geographic pattern of variation that is found similar by microsatellite markers (Lavery *et al.* 1996)

Population structure estimates conducted for males and females separately were very similar ($\Phi_{ST\ males}$ = 0.3478 or 0.3593 without CO; $\Phi_{STfemales}$ = 0.3537) and did not suggest male-biased dispersal. However, in order to detect male-biased dispersal by comparing mtDNA population structures estimates between sexes the bias needs to be strong and males need to disperse long distances, as haplotype similarity between nearby areas will obscure the extent of male dispersal. Strong male-biased dispersal was only evident between the adjacent populations of CL-FL (Table 5.4); therefore, it can be concluded that male-biased dispersal is weak in our study area. This result confirms previous findings of weak male-biased dispersal using microsatellite data for post-dispersal individuals (Chapter 4; Pérez-Espona *et al.* in prep.).

More insights into sex-specific dispersal patterns of Scottish red deer were given by the hierarchical AMOVAs conducted for each sex separately. Females showed lower levels of differentiation between regions than males which could be interpreted as historical gene flow between populations and/or current long distance dispersal by females. In contrast, males showed lower levels of genetic differentiation among populations within regions than females indicating that it is at small geographical scales at which male-biased dispersal occurs. Male-biased dispersal only at small geographical scales (< 5 km) was also found in previous analyses of post-dispersal individuals using microsatellite markers (Chapter 4; Pérez-Espona *et al.* in prep.) and average dispersal distances found in field-based studies (Daniels & McClean 2003; Sibbald *et al.* unpublished data). These findings are concordant with the idea that males do not need to move long distances to avoid mating with relatives as in a polygynous mammal the risk of inbreeding is lower than for females (Brooker *et al.* 1990). Furthermore, long distance dispersal of males might not be very beneficial as dispersal is risky and costly and individuals may arrive at new areas in worse body condition than resident males and be at disadvantage when competing for females (Johnson 1986). Therefore, staying in local or nearby areas waiting to acquire a

dominant rank might be a better strategy than dispersing to unknown new areas (Johnson 1986).

Red deer females are generally more philopatric than males and disperse shorter distances (Clutton-Brock *et al.* 1982a, 1989; Daniels & McClean 2003; Nussey *et al.*, 2005). However, in areas where competition for resources is important for reproduction, such as good quality grazing, competition between females is strong and females might be prone to disperse (Clutton-Brock *et al.* 1982b) and would benefit by dispersing longer distances in order to find better home ranges. Furthermore, field-based studies might not be able to distinguish among natal dispersal movements and dispersal during the mating season when males have been found to disperse larger distances than females (Sibbald *et al.* unpublished data). Although dispersal will be costly, the possibilities for reproduction in new areas will be much higher in females than for males (Johnson 1986). Therefore, longer distance colonisation events by females would be more effective than by males as females have a higher probability of passing their genes to the next generation even in new areas.

5.7 Conclusions

This is the most intense study ever conducted using mitochondrial control region sequences to assess genetic diversity and population structure of a mammal at a regional scale. The study has revealed important insights about the genetic diversity, population structure and gender specific dispersal patterns of Scottish Highland red deer, the largest population of red deer in Europe.

Despite several introduction events of foreign species of deer in Scotland, none of the Scottish Highland red deer included in this study presented introgressed mtDNA from other species or subspecies of deer. The influence of humans on red deer populations was evident in terms of translocations of individuals between British populations as a few haplotypes found in the Scottish mainland clustered very closely to haplotypes found in Euston (England) and some of the islands off the West

coast of Scotland where populations are known to partly or entirely have been originated from introduction events. Nonetheless, the minimum spanning network representing haplotype relationships and population structure analyses suggested that past management practices of Scottish Highland red deer have not completely blurred population boundaries as significant genetic structure was found between regions and population within regions, and private haplotypes were found in some localities.

Mismatch distribution analysis and neutrality tests suggested a population expansion for Scottish Highland red deer probably during the Pleistocene. Using an approximate mutation rate for the mtDNA control region the date for the population expansion was 93,000 years BP. However, as red deer are known to have been absent in Scotland during the last glaciations, only recolonising Scotland after the ice sheets retreated (c. 11,000 years BP) further studies are needed in order to provide an accurate date for the population expansion.

Comparisons between mtDNA control region and microsatellite pair-wise population differentiation measures indicated that landscape features along the Great Glen acting as red deer gene flow barriers are of a contemporary nature. Additionally, comparisons between genetic markers and mtDNA estimates of population structure between sexes indicated that male-biased dispersal is weak and probably occurs between nearby populations and that long distance gene flow between regions is more likely to be female-biased.

5.8 Acknowledgements

All stalkers and deer managers from the estates of Ardgour, Ardverikie, Ben Alder, Clunes, Conaglen, Corrour, Forest Lodge, Glencoe, Glencreran, Glenkinglass, Glenstrae, Kintail, Mamore and South Affric are greatly thanked for the collection of samples. Angela Sibbald and Russell Hooper are thanked for information on unpublished data from their GPS-based studies. Andy Gill and Jill Lovell are thanked

for their sequencing services at IEB. Kirsten Frew is thanked for help with some of the sequencing during her Honours project. This project was funded by The Macaulay Development Trust. SEERAD - Scottish Executive supported F.J.P.-B.

5.9 References

Abernethy K (1994) The establishment of a hybrid zone between red and sika deer (genus *Cervus*). *Molecular Ecology* **3**, 551-562.

Avise JC (1994) *Molecular Markers, Natural History and Evolution* Chapman & Hall, New York.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology* **13**, 729-744.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**, 155-165.

Boyce WM, Ramey II RR, Rodwell TC, Rubin ES, Singer RS (1999) Population subdivision among desert bighorn sheep (*Ovis canadensis*) ewes revealed by mitochondrial DNA analysis. *Molecular Ecology* **8**, 99-106. Brooker MG, Rowley MA, Adams M, Baverstock PR (1990) Promiscuity: an inbreeding avoidance mechanism in a socially monogamous species? *Behavioral Ecology and Sociobiology* **26**, 191-199.

Cathey JC, Bickham JW, Patton JC (1998) Introgressive hybridization and nonconcordant evolutionary history of maternal and paternal lineages in North American deer. *Evolution* **52**, 1224-1229.

Chesser RK, Baker RJ (1996) Effective sizes and dynamics of uniparentally and diparentally inherited markers. *Genetics* **144**, 1225-1235.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.

Clutton-Brock TH, Albon SD, Guinness FE (1982a) Competition between female relatives in a matrilocal mammal. *Nature* **300**, 178-180.

Clutton-Brock TH, Guinness FE, Albon SD (1982b) *Behaviour and Ecology of Two Sexes* University of Chicago Press, Chicago.

Clutton-Brock TH, Albon SD (1989) *Red deer in the Highlands* BSP Professional Books, Oxford.

Clutton-Brock TH, Albon SD (1992) Trial and error in the Highlands. *Nature* **358**, 11-12.

Clutton-Brock TH, Coulson TN, Milner AD, Thomson D, Armstrong HM (2002) Sex differences in emigration and mortality affect optimal management of deer populations *Nature* **415**, 633-637.

Clutton-Brock TH, Coulson T, Milner JM (2004) Red deer stocks in the Highlands of Scotland. *Nature* **429**, 261-262.

Crandall KA, Templeton AR (1993) Empirical tests of some predictions from coalescent theory with application to intraspecific phylogeny reconstruction. *Genetics* **134**, 959-969.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution* **15**, 290-295.

Daniels M, McClean C (2003) Red deer calf tagging programmes in Scotland - an analysis. *Deer (The Journal of the British Deer Society)* **12**, 420-423.

Donnelly P, Tavaré S (1986) The ages of alleles and a coalescent. *Advances in Applied probability* **18**, 1-19.

Edgar RC, (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797.

Eizirik E, Kim J-H, Menotti-Raymond M *et al.* (2001) Phylogeography, population history and conservatio genetics of jaguars (*Panthera onca*, Mammalia, Felidae).

Endler JA (1977) *Geographic Variation, Speciation, and Clines* Princeton University Press, Princeton, NJ.

Ennos RA (1994) Estimationg the relative rates of pollen and seed migration among plant populations. *Heredity* **72**, 250-259.

Escorza-Treviño S, Dizon AE (2000) Phylogeography, intraspecific structure and sex-biased dispersal of Dall's porpoise, *Phocoenoides dalli*, revealed by mitochondrial DNA analyses. *Molecular Ecology* **9**, 1049-1060.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **4**, 783-791.

Felsenstein J (2005) PHYLIP (Phylogeny Inference Package) Version 3.65. Distributed by the author. Department of Genome Sciences UoW, Seattle.

Feulner PGD, Bielfeldt W, Zachos FE (2004) Mitochondrial DNA and microsatellite analyses of the genetic status of the presumed subspecies *Cervus elaphus montanus* (Carpathian red deer). *Heredity* **93**, 299-306.

Filatov DA (2002) ProSeq: A software for preparation and evolutionary analysis of DNA sequence data sets. *Molecular Ecology Notes* **2**, 621-624.

Fitzsimmons NN, Moritz C, Limpus CJ, Pope L, Prince R (1997) Geographic structure of mitochondrial and nuclear gene polymorphisms in Australian green turtle populations and male-biased gene flow. *Genetics* **147**, 1843-1854.

Fu Y-X (1997) Statistical tests of neutrality of matations against population growth, hitchhiking and background selection. *Genetics* **147**, 915-925.

Fu Y-X, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* **133**, 693-709.

Goodman SJ, Barton NH, Swanson GM, Abernethy K, Pemberton JM (1999) Introgression through rare hybridization: A genetic study of a hybrid zone between red and sika deer (Genus *Cervus*) in Argyll, Scotland. *Genetics* **152**, 355-371.

Groves C (2006) The genus *Cervus* in eastern Asia. *European Journal of Wildlife Research* **52**, 14-22.

Harpending HC (1994) Signature of an ancient population growth in a low-resolution mitochondrial DNA mistmatch distribution. *Human Biology* **66**, 591-600.

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) Genetic structure of ancient human populations. *Current Anthropology* **34**, 483-496.

Hartl G, Zachos F, Nadlinger K (2003) Genetic diversity in European red deer (*Cervus elaphus* L.): anthropogenic influences on natural populations. *Comptes Rendus Biologies* **326**, 37-42.

Hartl GB, Zachos FE, Nadlinger K, *et al.* (2005) Allozyme and mitochondrial DNA analysis of French red deer (*Cervus elaphus*) populations: genetic structure and its implications for management and conservation. *Mammalian Biology* **70**, 24-34.

Hedrick PW (1999) Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution* **53**, 313-318.

Hedrick PW (2005) A standardised genetic differentiation measure. *Evolution* **59**, 1633-1638.

Hmwe SS, Zachos FE, Sale JB, Rose HR, Hartl GB (in press) Genetic variability and differentiation in red deer (*Cervus elaphus*) from Scotland and England. *Journal of Zoology.*

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 754-755.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77**, 282-291.

Johnson CN (1986) Sex-biased philopatry and dispersal in mammals. *Oecologia* **69**, 626-627.

Kasapidis P, Suchentrunk F, Magoulas A, Kotoulas G (2005) The shaping of mitochondrial DNA phylogeographic patterns of the brown hare (*Lepus europaeus*) under the combined influence of Late Pleistocene climatic fluctuations and anthropogenic translocations. *Molecular Phylogenetics and Evolution* **34**, 55-66.

Kerth G, Mayer F, Petit E (2002) Extreme sex-biased dispersal in the communally breeding, nonmigratory Bechstein's bat (*Myotis bechsteinii*) *Molecular Ecology* **11**, 1491-1498.

Ladoukakis ED, Zouros E (2001) Recombination in Animal Mitochondrial DNA: Evidence from Published Sequences. *Molecular Biology and Evolution* **18**, 2127-2131.

Lavery S, Moritz C, Fielder DR (1996) Genetic patterns suggest exponential population growth in a declining species. *Molecular Biology & Evolution* **13**, 1106-1113.

Lister AM (1984) Evolutionary and ecological origins of British deer. *Proceedings of the Royal Society of Edinburgh* **82B**, 205-299.

Liu XH, Wang YQ, Liu ZQ, Zu KY (2003) Phylogenetic relationships of Cervinae based on sequence of mitochondrial cytochrome b gene. *Zoological Research* **1**, 27-33.

Long JD (2003) *Introduced Mammals of the World: Their History, Distribution and Influence* CABI Publishing, Oxon.

Lowe VPW, Gardiner AS (1974) A re-examination of the subspecies of Red deer (*Cervus elaphus*) with particular reference to the stocks of Britain. *Journal of Zoology (London)* **174**, 185-201.

Ludt CJ, Schroeder W, Rottman O, Kuehn R (2004) Mitochondrial DNA phylogeography of red deer (*Cervus elaphus*). *Molecular Phylogenetics and Evolution* **3**, 1064-1083.

Lyrholm T, Leimar O, Johanneson B, Gyllensten U (1999) Sex-biased dispersal in sperm whales: contrasting mitochondrial and nuclear genetic structure of global populations. *Proceedings of the Royal Society, London* **266**, 347-354.

Maddison DR, Maddison WP (2000) McClade. Sinauer Associates, Sunderland, Massachusetts.

Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution* **9**, 373-375.

Moritz C (2002) Strategies to protect biological diversity. *Systematic Biology* **21**, 238-254.

Nagata J, Masuda R, Kaji K, Kaneko M, Yoshida MC (1998) Genetic variation and population structure of the Japanese sika deer (*Cervus nippon*) in Hokkaido Island, based on mitochondrial D-loop sequences. *Molecular Ecology* **7**, 871-877.

Nussey DH, Coltman DW, Coulson T, *et al.* (2005) Rapidly declining fine-scale spatial genetic structure in female red deer. *Molecular Ecology* **14**, 3395-3405.

Nussey DH, Pemberton J, Donald A, Kruuk LEB (in press) Genetic consequences of human management in an introduced island population of red deer (*Cervus elaphus*). *Heredity*.

Pereira F, Pereira L, Van Asch B, Bradley DG, Amorim A (2005) The mtDNA catalogue of all Portuguese autochthonous goat (*Capra hircus*) breeds: high diversity of female lineages at the western fringe of European distribution. *Molecular Ecology* **14**, 2313-2318.

Piertney SB, Maccoll AD, Bacon PJ, Racey PA, Lambin X, Dallas JF (2000). Matrilineal genetic structure and female-mediated gene flow in red grouse (*Lagopus lagopus scoticus*): an analysis using mitochondrial DNA. *Evolution* **54**, 279-289.

Pitra C, Fickel J, Meijaard E, Groves CP (2004) Evolution and phylogeny of old world deer. *Molecular Phylogenetics and Evolution* **33**, 880-895.

Polziehn RO, Hamr J, Mallory FF, Strobeck CC (1998) Phylogenetic status of North American wapiti (*Cervus elaphus*) subspecies. *Canadian Journal of Zoology* **76**, 998-1010.

Polziehn RO, Strobeck C (2002) A phylogenetic comparison of red deer and wapiti using mitochondrial DNA. *Molecular Phylogenetics and Evolution* **22**, 342-356.

Posada D, Crandall KA (1998) MODELTEST: Testing the model of DNA substitution. **14**, 817-818.

Prugnolle F, De Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity* **88**, 161-165.

Rambaut A, Drummond AJ (2004) TRACER [computer program]. Available at http://evolve.zoo.ox.ac.uk/software.html.

Ramey RR (1995) Mitochondrial DNA variation, population structure, and evolution of mountain sheep in the southwestern United States and Mexico. *Molecular Ecology* **4**, 429-439.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Journal of Moleculat Biology and Evolution* **19**.
Randi E, Mucci N, Claro-Hergueta F, Bonnet A, Douzery EJP (2001) A mitochondrial DNA control region phylogeny of the Cervinae: speciation in *Cervus* and implications for conservation. *Animal Conservation* **4**, 1-11.
Rice WR (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223-225.
Rodríguez F, Oliver JF, Marín A, Medina JR (1990) The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology* **142**, 485-501.
Rogers AR (1995) Genetic evidence for a Pleistocene population explosion. *Evolution* **49**, 608-615.
Rogers AR, Harpending HC (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**, 552–569.
Ronquist F, Huelsenbeck JP (2003) MRBAYES3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.
Ryman N, Jorde PE (2001) Statistical power when testing for genetic differentiation. *Molecular Ecology* **10**, 2361-2374.
Sarno RJ, Villalba L, Bonacic C, *et al.* (2004) Phylogeography and subspecies assessment of vicuñas in Chile and Bolivia utilizing mtDNA and microsatellite markers: implications for vicuña conservation and management. *Conservation Genetics* **5**, 89-102.
Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. *Genetics* **152**, 1079-1089.
Schneider S, Roessli D, Excoffier L (2000) Arlequin version 2.000: A software for population genetic analysis. Distributed by Genetics and Biometry Laboratory, University of Geneva, Switzerland.
Shaw CN, Wilson PJ, White BN (2003) A reliable molecular method of gender determination for mammals. *Journal of Mammalogy* **84**, 123-128.
Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences ins stable and exponentially growing populations. *Genetics* **129**, 555-562.
Tajima F (1989) The effect of change in population size on DNA polymorphism. *Genetics* **123**, 597-601.
Takahata N (1998) The coalescent in two partially isolated diffusion populations. *Genetics Research* **52**.
Tiedemann R, Hardy O, Vekemans X, Milinkovitch MC (2000) Higher impact of female than male migration on population structure in large mammals. *Molecular Ecology* **9**, 1159-1163.
Thulin CG, Stone J, Tegelstrom H, Walker CW (2006) Species assignment and hybrid identification among Scandinavian hares *Lepus europaeus* and *L. timidus*. *Wildlife Biology* **12**, 29-38.
Vernesi C, Pecchioli E, Caramelli D, Tiedemann R, Randi E, Bertorelle G (2002) The genetic structure of natural and reintroduced roe deer (*Capreolus capreolus*) populations in the Alps and central Italy, with reference to the mitochondrial DNA phylogeography of Europe. *Molecular Ecology* **11**, 1285-1297.

Whitehead GK (1960) *The deer stalking grounds of Great Britain and Ireland* Hollis and Carter, London.

Whitehead GK (1964) *The deer of Great Britain and Ireland* Routledge & Kegan Paul, London.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323–354.

Zachos FE, Hartl DL, Apollonio M, Reutershan T (2003) On the phylogeographic origin of the Corsican red deer (*Cervus elaphus corsicanus*): evidence from microsatellites and mitochondrial DNA. *Mammalian Biology* **68**, 284-298.

Chapter 6

Conclusions

6.1. Conclusions and implications for management of red deer in the mainland of Scotland

An understanding of the genetic diversity and population structure of red deer is crucial in order to provide a sustainable management plan for red deer on the mainland of Scotland. Knowledge of population boundaries is important in order to define the appropriate scale for management and identify management units (Moritz 1994). The effective management of a particular population of red deer might need the collaboration of the group of estates sharing the same population of red deer. For example, in estates with high densities of red deer where a reduction in numbers might be sought, culling only in one particular estate might not prove to be an effective strategy due to dispersal of individuals from neighbouring estates that might not be controlling population numbers. In addition, the delimitation of population boundaries and knowledge of the levels of connectivity between different subpopulations of red deer might also prove very valuable during disease outbreaks in order to control disease spread.

The different data chapters of this thesis provided important insights into the genetic diversity, population structure and some of the factors influencing the distribution of genetic diversity of red deer on the Scottish mainland; information that is not only important in order to understand the evolution and ecology of this native species but that would prove valuable if implemented in future management policies of Scottish red deer.

Chapter 2 described an efficient and reliable microsatellite genotyping multiplex high-throughput protocol which can not only be useful for future population genetic studies of red deer but has also the potential to be a powerful tool for parentage

analyses studies or for more applied studies such as forensic analyses of deer and detection of illegal trophy hunting.

Despite the relatively small scale of the study and the high dispersal capabilities of red deer, significant population genetic structure was found for red deer on the Scottish mainland; implying that the range of red deer on mainland Scotland is divided into different subpopulations interconnected by varying levels of gene flow. Adjacent or nearby estates seemed to share the same subpopulations or herds of red deer, as they were not significantly differentiated from each other. Therefore, management at local scale could benefit from collaboration between neighbouring estates that might otherwise considered as different management units. Although isolation-by-distance was significant across the study area, a strong gene flow barrier was found along the Great Glen. Analyses of the least-cost distance landscape matrices constructed using GIS techniques showed that sea lochs, mountain slopes, roads and fenced forests located in the Great Glen played an important role in the population differentiation observed. Furthermore, the assessment of a wide range of cost values used to calculate least-cost distance matrices provided an innovative approach on how to assign cost values for crossing a particular landscape cell and allowed the identification of further possible gene flow barriers and corridors.

Despite the fact that current management of red deer considers natural and man-made features when delimiting population boundaries, this study is the first to qualitatively and quantitatively assess the influence of landscape features on red deer population structure. Results obtained in Chapter 3 support the continuation of current delimitation of Deer Management Groups (DMGs) that consider natural and man-made features and the avoidance as far as possible of further divisions of DMGs, which might not represent natural population boundaries. Nearby estates separated by landscape features acting as barriers for red deer gene flow could despite the short geographical distance between estates, be holding different subpopulations of red deer and thus they might require different management strategies. In addition, knowledge of the effect of landscape features, in particular those man-made ones, on current red deer gene flow might also help to understand

how future human activities modifying the landscape could impact the genetic diversity and population structure of red deer in Scotland.

Sex-specific dispersal patterns of red deer in the Scottish mainland were shown not to strongly affect the population genetic structure. Sex-biased dispersal tests and spatial autocorrelation analyses comparing microsatellite data between post-dispersal male and female red deer indicated weak male-biased dispersal even at small geographical scales (< 0-5 km) (Chapter 4). Comparisons of population structure estimates obtained from microsatellites (Chapter 3) and mtDNA control region (Chapter 5) confirmed weak male-biased dispersal across the area and indicated that male-biased dispersal is likely to occur between neighbouring estates but that rarer longer dispersal events are more likely to be female-biased.

Analyses of mtDNA control region sequences revealed that past red deer management practices on the Scottish mainland, such as introductions of foreign deer species and translocations of red deer stocks, had a much more limited impact than it has had on the some of the Scottish islands, where part of or the entire population had originated from introduction events (Hmwe *et al.* in press; Nussey *et al.* in press). Despite the larger number of individuals sequenced in this study none presented introgressed mtDNA from other subspecies or species of deer, and only a very few individuals had haplotypes that might have originated from translocations of red deer into the Scottish mainland (Chapter 5). In addition, patterns of population structure estimated from mtDNA control region sequences were similar to those found with microsatellite markers, indicating that population boundaries of red deer have not been blurred despite the past management practices. This suggests that native Scottish red deer gene pools might still be present in the Scottish mainland.

Information on the levels of genetic diversity found within and among populations of Scottish mainland red deer and the identification of private mtDNA haplotypes in particular estates or groups of estates would also be valuable information to take into account when drawing future conservation and management programs in order to preserve Scottish red deer genetic diversity.

Insights into the population history of red deer in the mainland of Scotland were also provided by the distribution and relationships of mtDNA haplotypes, and by demographic analyses which suggested that current populations of red deer might have originated from an ancestral population that expanded from the Central Highlands into other parts of mainland Scotland. The estimated date of the expansion was found to be c. 93,000 years BP, which would fall within the last interglacial period in Western Europe (c. 135,000-70,000 years BP) when the red deer is known to have occurred in Britain through most or all the interglacial cycle (Lister 1984). Although accurate dating of the expansion requires further examination (see future research section), it nonetheless suggests that the expansion detected is not as a result of population fluctuations during the last few centuries but instead reflects more ancient demographic events.

Analyses of the mtDNA CR sequences also provided additional information on the effect of landscape features in red deer gene flow found in Chapter 3. The greatest population differentiation at mtDNA was found between geographically distant estates rather than among closely located populations at either side of the Great Glen as found using microsatellite data, suggesting that current rather than historical landscape features are responsible for the differentiation of red deer populations on either side of the Great Glen.

6.2. Future research

Past management practices involving the introduction of foreign subspecies or species of deer and translocations of different stocks of red deer into Scotland aimed at the improvement of trophy hunting quality and involved introductions of more males than females (Whitehead 1960, 1964). Due to the maternal inheritance of mitochondrial DNA, successful hybridisation between males of foreign subspecies or species of deer with red deer would remain undetectable. Research developing genetic markers in the male-specific Y chromosome showing variation between closely related species such as red deer, wapiti and sika deer will provide a tool with which to assess the extent of successful hybridisation and introgression in the male-

line. Y chromosome markers have already been used to trace male lineage introgression in other ungulates. For example, the Y chromosome marker ZFY was useful for identifying natural hybrids between the North American white-tailed deer (*Odocoileus virginianus*) and mule deer (*O. hemionus*) (Cathey *et al.* 1998).

Examination of mtDNA haplotypes for individuals from English deer parks from where introductions of female deer to Scotland are documented will help assess, in more detail, the success of past red deer stock translocations and to distinguish areas that are most likely to contained native Scottish red deer.

Molecular-clock based studies providing estimates for the mutation rate of the mtDNA CR of red deer will help to provide an accurate date for the population expansion of Scottish mainland red deer found in this study. An accurate date for the expansion will provide important insights into the evolution and population history of Scotland.

The large amount of genotypic data acquired during this study offers the possibility of a wide range of analyses examining different aspects of Scottish mainland red deer biology that due to time restrictions could not be covered in this thesis. For example, analyses examining the rate and direction of migration between populations and the identification of population sources and sinks, information that would be beneficial for management decisions. Analyses combining genetic data gathered during this study and ecological data for the populations analysed such as densities of red deer, sex ratios, quality of grazing and management practices conducted on individual estates would help explain current population structure and levels of gene flow between estates.

Analyses of the microsatellite genotypes obtained for 172 foetuses but not included in analyses conducted in Chapter 3, 4 and 5 could be used to provide a measure of effective male dispersal and to obtain an indication of the level of polygyny in red deer in mainland Scotland.

Long-term field-based studies following red deer calves into adulthood would provide further insights into the causes underlying the weak male-biased dispersal in the study area. For example, the following of male calves into adulthood could examine the possible behaviour of males returning to local areas during the mating season and to determine if related males tend to disperse together. In addition, depending on the nature of the data gathered from the calf tagging programmes, further analyses could also examine if the distances recorded for males could be attributed to natal dispersal or rut dispersal.

This study was conducted on red deer inhabiting open hill areas. However, 27,000-50,000 red deer are thought to live in forested areas (Clutton-Brock & Albon 1989) and therefore, similar studies to the ones presented in this thesis of red deer in forested areas would provide a more complete picture of the genetic diversity and population structure of red deer in mainland Scotland.

6.3 References

Cathey JC, Bickham JW, Patton JC (1998) Introgressive hybridization and nonconcordant evolutionary history of maternal and paternal lineages in North American deer. *Evolution* 52, 1224-1229.

Clutton-Brock TH, Albon SD (1989) *Red deer in the Highlands* BSP Professional Books, Oxford.

Hmwe SS, Zachos FE, Sale JB, Rose HR, Hartl GB (in press) Genetic variability and differentiation in red deer (*Cervus elaphus*) from Scotland and England. *Journal of Zoology*.

Lister AM (1984) Evolutionary and ecological origins of British deer. *Proceedings of the Royal Society of Edinburgh* 82B, 205-299.

Moritz C (1994) Defining 'evolutionary significant units' for conservation. *Trends in Ecology and Evolution* 9, 373-375.

Nussey DH, Pemberton J, Donald A, Kruuk LEB (in press) Genetic consequences of human management in an introduced island population of red deer (*Cervus elaphus*). *Heredity*.

Whitehead GK (1960) *The deer stalking grounds of Great Britain and Ireland* Hollis and Carter, London.

Whitehead GK (1964) *The deer of Great Britain and Ireland* Routledge & Kegan Paul, London.